

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DP/Genar

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILML)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
Gallagherburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@ncl.nih.gov
Shevchenko, Y., Wehner, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stancir, S., Thomas, P.J., Tjongson, B.B., Touchman, J.W., Tsougen, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILML at: <http://image.llnl.gov>
Series: IRAL Plate: 33 Row: C Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Location/Qualifiers
1..1964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:23280 IMAGE:4637504"
/tissue_type="lung, large cell carcinoma"
/clone_id="N1H_MGC_18"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
277..1260
/codon_start=1
/product="Unknown (protein for MGC:23280)"
/protein_id="AAH15582.1"
/db_xref="gi:15990434"

CDS

BASE COUNT 397 a 545 c 626 g 396 t
ORIGIN
Query Match 84.3% Score 1635; DB 9; Length 1964;
Best Local Similarity 99.9% Pred. No. 0;
Matches 1885; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
254 GCGCCAAACAGCGGATGAGAAAGTACGCGCTGAGCTGCGCGCGGAGACCGCGC 313
253 GCGCCAAACAGCGGATGAGAAAGTACGCGCTGAGCTGCGCGCGGAGACCGCGC 312
314 GTGTGTGCTGCTGCGGAGGAGGAGGCGGCGGCGCTGCTGAGCTCGCGAG 373
313 GTGTGTGCTGCTGCGGAGGAGGAGGCGGCGGCGCTGCTGAGCTCGCGAG 372
374 GAGAGTGGGAACATAGAGTATCTTCAATGAGCTTGAAGCTTGGCAATCTGAGT 433
373 GAGAGTGGGAACATAGAGTATCTTCAATGAGCTTGAAGCTTGGCAATCTGAGT 432
434 GCGGCTTTGCGAGCTCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 493
433 GCGGCTTTGCGAGCTCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 492
494 GCGGCTTTGCGAGCTCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 553

493 GCGGCTTTGCGAGCTCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 552
554 CATATCGGCTCTTCTGCTGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
553 CATATCGGCTCTTCTGCTGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
614 AGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
613 AGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
674 CTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 733
673 CTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732
734 GCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
733 GCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
794 TATGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853
793 TATGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852
854 CTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
853 CTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
914 CAGACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 973
913 CAGACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
974 GCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1033
973 GCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1032
1034 CTATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1093
1033 CTATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092
1094 GAAGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1153
1093 GAAGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1152
1154 GAGCCCAAGCTTCTCAACCTTACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1213
1153 GAGCCCAAGCTTCTCAACCTTACCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1212
1214 ACGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273
1213 ACGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
1274 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1333
1273 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332
1334 TGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1393
1333 TGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
1394 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1453
1393 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1452
1454 GACAGTACCCAGATGAGATAGAGGATAGAGGATAGAGGATAGAGGATAGAGG 1513
1453 GACAGTACCCAGATGAGATAGAGGATAGAGGATAGAGGATAGAGGATAGAGG 1512
1514 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1573
1513 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
1574 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1633

[illegible][illegible]

repeat_region /rpt family="AluSx"
complement(17920..8220)
repeat_region /rpt family="AluSx"
complement(8364...8662)
repeat_region /rpt family="AluY"
8694...8831
repeat_region /rpt family="AluUb"
8832...9133
repeat_region /rpt family="AluSg"
9143...9415
repeat_region /rpt family="AluSg"
10270...10279
repeat_region /rpt family="AluSg"
10280...10574
repeat_region /rpt family="AluSx"
10575...10590
repeat_region /rpt family="AluSg"
10591...10872
repeat_region /rpt family="AluSg"
10873...10886
repeat_region /rpt family="AluSg"
11232...11395
repeat_region /rpt family="AluSg"
11396...11711
repeat_region /rpt family="AluY"
11712...11728
repeat_region /rpt family="AluSg"
12424...12558
repeat_region /rpt family="AluSg"
complement(13786...14088)
repeat_region /rpt family="AluSx"
complement(14100...14401)
repeat_region /rpt family="AluSg"
complement(14524...14691)
repeat_region /rpt family="AluSg"
14827...15139
repeat_region /rpt family="AluSx"
15191...15283
repeat_region /rpt family="AluSg"
complement(15224...15467)
repeat_region /rpt family="AluSg"
15496...15781
repeat_region /rpt family="AluSg"
15788...15981
repeat_region /rpt family="AluSg/x"
complement(15983...16040)
repeat_region /rpt family="AluSg/x"
16055...16220
repeat_region /rpt family="AluSg/x"
16330...16510
repeat_region /rpt family="AluSg/x"
complement(16512...16745)
repeat_region /rpt family="AluSg/x"
16746...17024
repeat_region /rpt family="AluSg/x"
complement(17025...17325)
repeat_region /rpt family="AluSg/x"
complement(17331...17625)
repeat_region /rpt family="AluSg/x"
complement(17455...17459)
repeat_region /rpt family="AluSg/x"
complement(17525...17538)
repeat_region /rpt family="AluSg/x"
complement(17549...17588)
repeat_region /rpt family="AluSg/x"
complement(17599...17607)
repeat_region /rpt family="AluSg/x"
complement(17623...17627)
repeat_region /rpt family="AluSg/x"
complement(17932...18209)
repeat_region /rpt family="AluSg/x"

Query Match 54.7% Score 1061, DB 9, Length 175020;
Best Local Similarity 99.9% Pred. No. 0;
Matches 1111, Conservative 0, Mismatches 1, Indels 0, Gaps 0;
QY CAGGCGCTGTGAACCTGGAAGCTGTTCCTGCGCAATGTTCTGTGAATGCTGCGCCCACTT 866
DB CAGGCGCTGTGAACCTGGAAGCTGTTCCTGCGCAATGTTCTGTGAATGCTGCGCCCACTT 91984
QY TCGCCCAATGAGCTTGGCTGGTGTCCGGGCAACGAAGGGGGTCCCGAGACCCCTGT 926
DB TCGCCCAATGAGCTTGGCTGGTGTCCGGGCAACGAAGGGGGTCCCGAGACCCCTGT 91924
QY ATTGTGCTCTACAGAGGAGCATGAGCCCTCTAGTGGAGATATTTGGCACTGGCATG 986
DB ATTGTGCTCTACAGAGGAGCATGAGCCCTCTAGTGGAGATATTTGGCACTGGCATG 91864
QY TGAAGAGCTGCTTCACTGCTCCGAGACGACCCGGGCAACCCATGCGCTATGGAGCCA 1046
DB TGAAGAGCTGCTTCACTGCTCCGAGACGACCCGGGCAACCCATGCGCTATGGAGCCA 91804
QY GCAAGAGCTGAGAGGCTTGGGCTGGGAGAGATGTAACCCGATGAAGACCCCATG 1106
DB GCAAGAGCTGAGAGGCTTGGGCTGGGAGAGATGTAACCCGATGAAGACCCCATG 91744
QY CTGAGAGCTCAGAGGCCCCATCTTCTTAAGACACCCCACTGAGAGAGCCCAAGTT 1166
DB CTGAGAGCTCAGAGGCCCCATCTTCTTAAGACACCCCACTGAGAGAGCCCAAGTT 91684
QY CTCAACCTTACCCCAAGCCCTCAGAGCTCAACAGATTTGTCTAAGATGACGACCCGATTC 1226
DB CTCAACCTTACCCCAAGCCCTCAGAGCTCAACAGATTTGTCTAAGATGACGACCCGATTC 91624
QY AGGCTAAGTGAAGCTGAGATCCAGCTCTCTTAACCCCTGAGAGAGATGCTGCGCATG 1286
DB AGGCTAAGTGAAGCTGAGATCCAGCTCTCTTAACCCCTGAGAGAGATGCTGCGCATG 91624
QY GCACTTCAATGCTCTTGAAGAACTCTGAGATGCTGAGAGAGAGAGAGAGAGAGAGAG 1346
DB GCACTTCAATGCTCTTGAAGAACTCTGAGATGCTGAGAGAGAGAGAGAGAGAGAGAG 91504
QY GTTGTGATCTTGAACCTCCGCTGTTAATTTCTGAGAGAGAGAGAGAGAGAGAGAG 1406
DB GTTGTGATCTTGAACCTCCGCTGTTAATTTCTGAGAGAGAGAGAGAGAGAGAGAG 91444
QY CTTTCCGTTGAAGAGATATGAGGATATTTCTTCTGAGAGAGAGAGAGAGAGAGAGAG 1466
DB CTTTCCGTTGAAGAGATATGAGGATATTTCTTCTGAGAGAGAGAGAGAGAGAGAGAG 91384
QY GATGAGAGATGAGGAGATGCTAGACCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAG 1526
DB GATGAGAGATGAGGAGATGCTAGACCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAG 91324
QY CAGGCCCCCACTTATGATCTGATCAGCTCTGAGAGAGAGAGAGAGAGAGAGAG 1586
DB CAGGCCCCCACTTATGATCTGATCAGCTCTGAGAGAGAGAGAGAGAGAGAGAG 91264
QY GATGAG 1646
DB GATGAG 91204
QY GTTAATTAACCCCAATGTTAATGAAGCGAATTAAGGCTCCGAGCTAAGAGAGAGAG 1706
DB GTTAATTAACCCCAATGTTAATGAAGCGAATTAAGGCTCCGAGCTAAGAGAGAGAG 91144
QY GGGTCTCACTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1766
DB GGGTCTCACTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 91084
QY CGAGCTGCGTAAGATGAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1826
DB CGAGCTGCGTAAGATGAGTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 91024

QY 1827 GCCCATGAGAGTAAGGAGACGCTTCCGGGCGGATGCAAGGCTGGGTGATCTGATCT 1886
DB 91023 GCCCATGAGAGTAAGGAGACGCTTCCGGGCGGATGCAAGGCTGGGTGATCTGATCT 90964
QY 1887 GAAGCCCTCGAGATAAAGCGCTTGAACGCC 1918
DB 90963 GAAGCCCTCGAGATAAAGCGCTTGAACGCC 90932

RESULT 4
AX071778 397 bp. DNA linear PAT 25-JAN-2001
LOCUS Sequence 2250 from Patent WO0102568.
ACCESSION AX071778
VERSION AX071778.1 GI:12582129
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,
Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
Lamson, G., Drmanac, R., Crkvenjakov, R., Drmanac, S., Dickson, M.,
Luhac, I., Leskowitz, D., Kita, D., Garcia, V., and Strache-Crain, B.
Human genes and gene expression products
Patent: WO 0102568-A 2250 11-JAN-2001;
CHIRON CORPORATION (US) / HYSEQ, INC. (US)

TITLE
JOURNAL
FEATURES
SOURCE 1. 397
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 74 a 116 c 125 g 82 t

ORIGIN
Query Match 10.3%; Score 200; DB 6; Length 397;
Best Local Similarity 100.0%; Pred. No. 3.3e-97;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 716 TATGCTGACATGAGTGGCTGATGATGCTGTTGCCCGGAGCTGGCAACACACTTGG 775
DB 31 TATGCTGACATGAGTGGCTGATGATGCTGTTGCCCGGAGCTGGCAACACACTTGG 90
QY 776 GCCACTGGCGTACCTGCTATGACGCCACCCAGGCGCTGTGAATCGAGCTGTTCTG 835
DB 91 GCCACTGGCGTACCTGCTATGACGCCACCCAGGCGCTGTGAATCGAGCTGTTCTG 150
QY 836 CGGCATGTTCTGATGAGTGGCTGGCCCACTTTGGCCCATTTGGCTGCTGCTCGG 895
DB 151 CGGCATGTTCTGATGAGTGGCTGGCCCACTTTGGCCCATTTGGCTGCTGCTCGG 210
QY 896 GCACCAAGAGGGGGTGGCCCA 915
DB 211 GCACCAAGAGGGGGTGGCCCA 230

RESULT 5
AL669840 10355 bp. DNA linear ROD 05-APR-2002
LOCUS Mouse DNA sequence from clone RP23-43J1 on chromosome 11, complete
sequence.
ACCESSION AL669840
VERSION AL669840.5 GI:20068701
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus (house mouse)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Bates, K.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,

COMMENT
Cambridgehire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On Apr 7, 2002 this sequence version replaced gi:19309585.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WormPeP; Information on the WormPeP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-43J1 is
from the RP23-43J1 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES
SOURCE 1. 103555
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-43J1"
/clone_1b="RP23-23"

BASE COUNT 27308 a 23935 c 24282 g 28026 t

ORIGIN
Query Match 2.7%; Score 52; DB 10; Length 10355;
Best Local Similarity 100.0%; Pred. No. 1.3e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 284 GCGCTGAGAGTGGCGGCGGAGGCGCGGTGCTGCTGCGCGGACGCC 335
DB 17422 GCGCTGAGAGTGGCGGCGGAGGCGCGGTGCTGCTGCGCGGACGCC 17473

RESULT 6
AC048361 204628 bp. DNA linear HTG 27-MAR-2003
LOCUS Mus musculus chromosome 11 clone RP23-163A8 map 11, WORKING DRAFT
SEQUENCE, 8 unordered pieces.
ACCESSION AC048361
VERSION AC048361.4 GI:29294161
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus (house mouse)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Birren, B., Nuebaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barne, N., Baerlein, V., Bada, F.,
Bogunlavsky, L., Boukhalter, B., Brown, A., Burkett, G.,
Campopiano, A., Casale, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodgson, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galegan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karitas, A.,
Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehotzky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Margulis, N.,

TITLE
JOURNAL

REFERENCE
AUTHORS

McCarthy, M., McEwen, P., McGurk, A., McKernan, K., McSheeters, R., Meltrin, J., Menes, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talmas, J., Testaye, S., Theodore, J., Tittel, A., Travers, M., Tiggililo, J., Vassilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Ye, W., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (14-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 204628)

TITLE
JOURNAL

REFERENCE
AUTHORS

Birren, B., Nussbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, S., Arachchi, H.M., Batra, N., Buetler, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarero, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, J., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, J., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Gage, J., Gargyala, S., Gerhart, J., Grand-Pierre, N., Hales, N., Hargopian, D., Hargos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Nordu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunphang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rice, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Talmas, J., Testaye, S., Theodore, J., Tittel, A., Travers, M., Tiggililo, J., Vassilev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (12-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 27, 2003 this sequence version replaced gi:8076973.
All repeats were identified using RepeatMasker:
Smith, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

Project Information

Center project name: 163_A_8

Center clone name: 163_A_8

Summary Statistics

Sequencing vector: M13; M77815; 3% of reads

Sequencing vector: plasmid; n/a; 67% of reads

Chemistry: Dye-terminator Big Dye 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 202765 bases at least Q40

Consensus quality: 203217 bases at least Q30

Consensus quality: 203451 bases at least Q20

Insert size: 19100; agarose-efp

Insert size: 203928; sum-of-contigs

Quality coverage: 12.1 in Q20 bases; agarose-efp

Quality coverage: 11.3 in Q20.

NOTE: This is a 'working draft' sequence. It currently

consists of 8 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 81112: contig of 81112 bp in length

81113 81112: gap of 100 bp

81213 90109: contig of 8897 bp in length

90209: gap of 100 bp

FEATURES
source

90210 102935: contig of 12726 bp in length

102936 103035: gap of 100 bp

103036 116669: contig of 13634 bp in length

116670 116769: gap of 100 bp

116770 131956: contig of 15187 bp in length

131957 131957: gap of 100 bp

132057 132056: gap of 100 bp

132057 148455: contig of 16399 bp in length

148456 148555: gap of 100 bp

148556 200359: contig of 51804 bp in length

200360 200459: gap of 100 bp

200460 204628: contig of 4169 bp in length.

Location/Qualifiers

1..204628

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/chromosome="11"

/map="11"

/clone="RP23-163A8"

/clone_11b="RP21-23 Female Mouse BAC"

1..81112

/note="assembly_fragment"

/clone_end:SP6

/vector_side:left

81213..90109

/note="assembly_fragment"

90210..102935

/note="assembly_fragment"

103036..116669

/note="assembly_fragment"

116770..131956

/note="assembly_fragment"

132057..148455

/note="assembly_fragment"

148556..200359

/note="assembly_fragment"

200460..204628

/note="assembly_fragment"

/clone_end:TV

/vector_side:right

BASE COUNT 54328 a 48238 c 47410 g 53947 t 705 others

ORIGIN

Query Match 2.7% Score 52; DB 2; Length 204628;

Best Local Similarity 100.0%; Pred. No. 1.2e-16;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

244 GCGCTGAGAGCTGGCGCGCGCGCGCGCGCGCGCGCTGCTGCTGCGCGCGCGCGC 335

DB 168980 GCGCTGAGAGCTGGCGCGCGCGCGCGCGCGCGCGCTGCTGCTGCGCGCGCGC 169031

RESULT 7

AX163165 51 bp DNA linear PAT 22-JUN-2001

LOCUS AX163165 Sequence 6493 from Patent WO0140521.

DEFINITION AX163165

ACCESSION AX163165

VERSION AX163165.1 GI:14544496

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

Shinkens, R.A. and Leach, M.

Nucleic acids containing single nucleotide polymorphisms and

methods of use thereof

Patent: WO 0140521-A 6493 07-JUN-2001;

Curegen Corporation (US)

Location/Qualifiers

1..51

/organism="Homo sapiens"

/mol_type="genomic DNA"

```

misc_feature      /db_xref="taxon:9606"
                  26
                  /note="1 of 2 allelic variants (6494 is other entry)
Accession number CG34087949"
BASE COUNT      6 a 17 c 18 g 10 t
ORIGIN
Query Match      2.6% Score 51; DB 6; Length 51;
Beet Local Similarity 100.0%; Prid. No. 6e-16;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 594 GCCTGAAGGACATGTGCCTTACGCGCGGTGGTGTGAAGCTCAGACTGCC 644
1 GCCTGAAGGACATGTGCCTTACGCGCGGTGGTGTGAAGCTCAGACTGCC 51

Db 1 GCCTGAAGGACATGTGCCTTACGCGCGGTGGTGTGAAGCTCAGACTGCC 51

RESULT 8
AC114124
LOCUS
DEFINITION
AC114124 245259 bp DNA linear HTG 13-MAY-2003
Rattus norvegicus clone CH230-40F15, *** SEQUENCING IN PROGRESS
***, 2 unordered pieces.
AC114124
AC114124.4 GI:30579659
HTG, HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 245259)
Muzny, D.Marie, Metzker, M.Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alibrooke, S., Amin, A., Angiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, I.M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biewald, K., Blair, J., Blankenburg, K., Bjyht, P., Brown, M.,
Bryant, N., Bubay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
Chacko, J., Chavez, R., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
David, L.M.L., Davis, C., Davy-Carroll, L., De Ande, C., Dedrich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dint, H., Dlyva, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durkin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C.A., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Galla, R., Garcia, A., Garner, T., Garza, M.,
Georgescu, J.B., Geer, K., Giller, R., Grady, M., Guerra, W., Guevara, M.,
Gunarekne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K.,
Harvey, J., Havlik, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulik, S., Humm, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpaty, S., Kelly, S., Kelly, S., Khan, Z., Kling, L., Kovac, C.,
Kovacs, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshaw, L., Louissege, H., Lozado, R.O., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mammond, M., Malloy, K., Mangun, A.,
Mangum, B., Mapua, P., Martin, K., Martin, K., Martinez, B.,
Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
Milosavljevic, A., Miner, G., Mnja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L.,
Nandeyia, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwokweli, O., Okwomou, G., Olampunsaagoun, A., Pal, S., Parks, K.,
Paeterek, S., Paul, H., Perez, A., Perez, L., Pfankoch, C.,
Plopper, F., Polindexter, A., Popovic, D., Prims, B., Pu, L.,
Puzos, M., Quiroz, J., Rachlin, E., Reeves, K., Reiter, M.A., Reigh, R.,
Rivley, B., Rellly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodery, T., Rojas, A., Rose, R., Rose, R., Ruiz, S.O.,
Sanders, C., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smaj, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Stelmle, M., Strong, R., Sutcon, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, R., Thomas, N., Thomas, S., Tingey, A., Tjof, Z., Umatt, K.,
Valdes, R., Vera, V., Villanana, D., Waltron, A., Walker, B., Wang, J.,

```

```

TITLE      JOURNAL
REFERENCE  AUTHORS
REFERENCE  TITLE
JOURNAL    JOURNAL

Direct Submission
Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 245259)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:23269093.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly ('a' contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GLND
Center clone name: CH230-40F15
Summary Statistics
Assembly program: Atlas 3.0j
Consensus quality: 195275 bases at least Q40
Consensus quality: 199584 bases at least Q30
Consensus quality: 202510 bases at least Q20
Estimated insert size: 205179; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 244085: contig of 244085 bp in length
* 244086 244185: gap of unknown length
* 244186 245259: contig of 1074 bp in length.
Location/Qualifiers
1. 245259
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-40F15"
14909..17613
/note="wgs contig"
35833..38580
/note="wgs contig"
91665..93428
misc_feature
misc_feature
misc_feature
misc_feature

```


/note="wgs_contig"
 mlec_feature 113391. 114894
 /note="wgs_contig"
 BASE COUNT 53423 a 48733 c 48928 g 53896 t 40279 others
 ORIGIN

Query Match 2.2%; Score 42; DB 2; Length 245259;
 Best Local Similarity 100.0%; Pred. No. 3.5e-13;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 383 AACATGAGTCATCTTCATGCGCTTCGACCTTGCGCAATCTG 424
 DB 133948 AACATGAGTCATCTTCATGCGCTTCGACCTTGCGCAATCTG 133989

RESULT 9
 AC133243 177877 bp DNA linear HTG 10-SEP-2002
 LOCUS Rattus norvegicus clone CH230-10014. ** SEQUENCING IN PROGRESS
 DEFINITION *** 59 unordered pieces.

ACCESSION AC133243
 VERSION AC133243.2 GI:22772764
 KEYWORDS HTG; HTGS_PHASE1.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 1 (bases 1 to 177877)
 Muzny, D., Marle, Metzger, M., Lee, A., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D.,
 Anyalelechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biscaldi, K., Blair, J., Blankensbury, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Cechlar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M. L., Davis, C., Day-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Gatta, R., Garcia, A., Garner, T., Garza, M.,
 Gegregoridis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hughes, M.,
 Hollins, B., Howells, S., Hulik, S., Hume, J., Idelchid, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenshew, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangun, A.,
 Mangum, B., Mapuz, P., Martin, K., Martin, K., Martinez, E.,
 Mawhinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A.,
 Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,
 Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
 Newton, N., Nguyen, N., Norris, S., Nwakoelameh, O., Okunolu, G.,
 Olariunasoon, A., Pal, S., Parks, K., Paternak, S., Paul, H.,
 Perez, A., Perez, L., Pflankoch, C., Plopper, F., Polidexter, A.,
 Popovic, D., Primus, E., Pu, L., Pu, L., Puzo, M., Quiroz, J., Rachlin, E.,
 Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Rem, Y.,
 Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A.,
 Rose, M., Rose, R., Ruiz, S. J., Sanders, M., Savery, G., Scherer, S.,
 Scott, G., Shattman, S., Shen, H., Shetty, J., Shvartsbeyn, A.,
 Sison, I., Sitter, C. D., Smajd, D., Sneed, A., Sodergren, E.,
 Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A.,
 Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, D.,
 Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villaseana, D.,
 Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J.,
 Warren, R., Wei, X., White, P., Williams, G., Wilson, R., Wlasczyk, R.,

JOURNAL
 TITLE
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
 Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, Y., Zhou, J., Zhou, X.,
 Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R.,
 Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.
 Direct Submission
 Unpublished
 2 (bases 1 to 177877)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (08-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 177877)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (10-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Sep 10, 2002 this sequence version replaced gi:22758715.

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GDRJ
 Center clone name: CH230-10014
 Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye 1000 of reads
 Assembly program: Phrap, version 0.990329
 Consensus quality: 120753 bases at least Q40
 Consensus quality: 127137 bases at least Q30
 Consensus quality: 132199 bases at least Q20

NOT: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_direct_data.html).
 NOTE: This is a 'working draft' sequence. It currently
 consists of 59 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

1	1033	contig of 1033 bp in length
1034	1133	gap of unknown length
1134	2150	contig of 1017 bp in length
2151	2250	gap of unknown length
2251	3283	contig of 1033 bp in length
3284	3384	gap of unknown length
3384	4384	contig of 1001 bp in length
4385	4484	gap of unknown length
4485	5726	contig of 1242 bp in length
5727	5826	gap of unknown length
5827	6841	contig of 1015 bp in length
6842	6941	gap of unknown length
6942	8340	contig of 1399 bp in length
8341	8440	gap of unknown length
8441	9752	contig of 1312 bp in length
9753	9852	gap of unknown length
9853	11013	contig of 1161 bp in length
11014	11114	gap of unknown length
11114	12508	contig of 1395 bp in length
12509	12608	gap of unknown length
12609	14066	contig of 1458 bp in length
14067	14166	gap of unknown length
14167	15223	contig of 1057 bp in length
15224	15323	gap of unknown length
15324	16817	contig of 1494 bp in length
16818	16917	gap of unknown length
16918	18907	contig of 1990 bp in length
18908	19007	gap of unknown length

Query Match

Query Match	Score 31, DB 2, Length 177877;
19008	20556: contig of 1549 bp in length
20557	20656: gap of unknown length
20657	22042: contig of 136 bp in length
22043	22142: gap of unknown length
22143	23173: contig of 1031 bp in length
23174	23273: gap of unknown length
23274	24424: contig of 1151 bp in length
24425	24524: gap of unknown length
24525	26395: contig of 1871 bp in length
26396	26496: gap of unknown length
26496	27656: contig of 1161 bp in length
27657	28988: gap of unknown length
28988	29088: gap of unknown length
29089	31014: contig of 1926 bp in length
31015	31114: gap of unknown length
31115	33214: contig of 2100 bp in length
33215	33314: gap of unknown length
33315	35513: contig of 2199 bp in length
35514	35613: gap of unknown length
35614	37742: contig of 2129 bp in length
37743	37842: gap of unknown length
37843	38950: contig of 1108 bp in length
38951	39050: gap of unknown length
39051	41722: contig of 2672 bp in length
41723	41822: gap of unknown length
41823	43959: contig of 2137 bp in length
43960	44059: gap of unknown length
44060	47505: contig of 3446 bp in length
47506	47605: gap of unknown length
47606	50441: contig of 2836 bp in length
50442	50541: gap of unknown length
50542	53541: contig of 3000 bp in length
53542	53641: gap of unknown length
53642	55665: contig of 2024 bp in length
55666	55765: gap of unknown length
55766	57645: contig of 1880 bp in length
57646	57745: gap of unknown length
57746	59602: contig of 1857 bp in length
59603	59702: gap of unknown length
59703	62147: contig of 2445 bp in length
62148	62247: gap of unknown length
62248	65121: contig of 2874 bp in length
65122	65221: gap of unknown length
65222	68430: contig of 3229 bp in length
68431	68550: gap of unknown length
68551	73413: contig of 4863 bp in length
73414	73513: gap of unknown length
73514	76578: contig of 1065 bp in length
76579	76678: gap of unknown length
76679	79293: gap of unknown length
79294	79393: contig of 2617 bp in length
79396	81954: gap of unknown length
81955	82054: contig of 2559 bp in length
82055	85014: contig of 2960 bp in length
85015	85114: gap of unknown length
85115	89088: contig of 3974 bp in length
89089	93407: gap of unknown length
93408	93507: gap of unknown length
93508	97240: contig of 3733 bp in length
97241	97340: gap of unknown length
97341	103099: contig of 5759 bp in length
103100	103199: gap of unknown length
103200	106048: contig of 2849 bp in length
106049	106148: gap of unknown length
106149	113083: contig of 5935 bp in length
113084	113183: gap of unknown length
113185	115595: contig of 3412 bp in length
115596	123308: contig of 6613 bp in length

Best Local Similarity 100.0%; Pred. No. 3,5e-05; Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1909 GTTGACCGCCCAAAAAAAAAAAAAAAAAA 1939

Db 94363 GTTGACCGCCCAAAAAAAAAAAAAAAAAA 94333

RESULT 10

AB054499 413 bp DNA linear MAM 27-JUN-2001

LOCUS Physeter catodon DNA, SINE flanking sequence Sp2 locus.

DEFINITION AB054499

ACCESSION AB054499.1 GI:114549337

VERSION

KEYWORDS

SOURCE

ORGANISM

Physeter catodon (sperm whale)

Physeter catodon

Bukayocra, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;

Physeteridae; Physeter.

REFERENCE

AUTHORS

1 (sites)

Nikaido, M., Matsuo, F., Hamilton, H., Brownell, R.L., Jr., Cao, Y.,

Ding, W., Zuo, Y., Schedlock, A.M., Fordyce, R.E., Hasegawa, M. and

Okada, N.

Retrospect analysis of major cetacean lineages: the monophyly of

toothed whales and the paraphyly of river dolphins

Proc. Natl. Acad. Sci. U.S.A. 98 (13), 7384-7389 (2001)

2310001

PUBMED

11416211

REFERENCE

AUTHORS

2 (bases 1 to 413)

Nikaido, M.

Direct Submission

Submitted (17-JAN-2001) Masato Nikaido, Tokyo Institute of

Technology, Department of Biological Sciences, Midori-ku

Nagatsuta-cho 4259, Yokohama-shi, Kanagawa-ken 2268501, Japan

(E-mail:mnikaido@bio.itech.ac.jp, Tel:81459245742,

Fax:8145924835)

FEATURES

source

1. 413

/organism="Physeter catodon"

/mol_type="genomic DNA"

/db_xref="taxon:9755"

misc_feature 1..413

/note="SINE flanking sequence Sp2 locus"

BASE COUNT 153 a 77 c 112 g 71 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 0.0068; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1913 ACCGCCCAAAAAAAAAAAAAAAAAA 1939

Db 310 ACCGCCCAAAAAAAAAAAAAAAAAA 336

RESULT 11

LOCUS G63447

DEFINITION SHGC-141356 Human Homo sapiens STS genomic, sequence tagged site.

ACCESSION G63447

VERSION G63447.1 GI:6600566

KEYWORDS

STS.

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Bukayocra, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 561)

Olivier, M. and Cox, D.R.

Unpublished, Olivier, M., Cox, D.R. (2000)

Unpublished (2000)

COMMENT

Contact: Michael Olivier, David R. Cox

Stanford Human Genome Center
Stanford University School of Medicine
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
Tel: (650) 320-5800
Fax: (650) 320-5801
Email: olivier@shgc.stanford.edu
Primer A: AATTCAGCTGTGATCTGAGGC
Primer B: TGATCCACGAGCTGAGAACGA
STS size: 346
PCR Profile:

Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9700
Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
AmpliTaq Gold Polymerase: 0.07 units/ul
Total Vol: 5 ul

Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

Finished human sequence in NCBI. STSs designed and developed at the
Stanford Human Genome Center.
Location/Qualifiers
1..561

FEATURES
source
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="12"
/clone_11b="Human"
STS
primer bind 208..553
primer bind complement(531..553)
BASE COUNT 176 a 101 c 133 g 151 t
ORIGIN

Query Match 1.4%; Score 27; DB 11; Length 561;
Best Local Similarity 100.0%; Pred.No. 0.0067;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1913 ACCGCCCAAAAAAAAAAAAAAAAAA 1939
DB 491 ACCGCCCAAAAAAAAAAAAAAAAAA 517

RESULT 12
LDAJ3162/c 616 bp DNA linear PLN 12-JAN-2001
LOCUS LDAJ3162
DEFINITION Laminaria digitata DNA, simple tandem repeat Ld2-158.
ACCESSION AJ003162
VERSION AJ003162.1 GI:3392938
KEYWORDS
SOURCE Laminaria digitata
ORGANISM Laminaria digitata
Eukaryota; stramenopiles; Phaeophyceae; Laminariales;
Laminariaceae; Laminaria.

REFERENCE
AUTHORS Billoot C., Rousvoal S., Bercou A., Epien J.T.,
Saunton-Leprade P., Valero M. and Klotz G.S.
TITLE Isolation and characterization of microsatellite markers in the
nuclear genome of the brown alga Laminaria digitata (Phaeophyceae)
JOURNAL Mol. Ecol. 7 (12), 1778-1780 (1998)
MEDLINE 99076304
PUBMED 9859207
REFERENCE 2 (bases 1 to 616)

AUTHORS Billoot C.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-1997) Billoot C., Biologie Cellulaire et
Moléculaire des Algues, Station Biologique CNRS, Place Georges
Telesier, ROSCOFF, F-29682, FRANCE
COMMENT uncoding region.
FEATURES location/Qualifiers
source 1..616
/organism="Laminaria digitata"
/mol_type="genomic DNA"
/db_xref="taxon:80365"
/clone_11b="bq2-158"

BASE COUNT 167 a 134 c 143 g 172 t
ORIGIN

Query Match 1.4%; Score 27; DB 8; Length 616;
Best Local Similarity 100.0%; Pred.No. 0.0066;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1913 ACCGCCCAAAAAAAAAAAAAAAAAA 1939
DB 590 ACCGCCCAAAAAAAAAAAAAAAAAA 564

RESULT 13
BC017773 905 bp mRNA linear PRI 06-DEC-2001
LOCUS BC017773
DEFINITION Homo sapiens, triggering receptor expressed on myeloid cells 1,
clone MGC:22242 IMAGE:4692680, mRNA, complete cds.
ACCESSION BC017773.1 GI:17389458
VERSION BC017773.1 GI:17389458
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 905)
Strausberg, R.
DIRECT SUBMISSION
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA

REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpax11.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 36 Row: m Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 8924261.
Location/Qualifiers
1..905
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:54210"
/db_xref="taxon:9606"
/clone="MGC:22242 IMAGE:4692680"
/tissue_type="lung"
/clone_11b="NIH MGC_77"
/lab_host="DH10B"
/note="vector: pDNR-LIB"

CDS

46..750

/codon_start=1
/product="triglycerin receptor expressed on myeloid cells
1"
/protein_id="AAH1773.1"
/db_xref="GI:17389459"
/translation="MRKRLMGILNMLFVSEI,RAATKLTREKYEKEGTLDPKCDY
LKKPSSOKAKMIDIRDEMPKTLACTERPSKSHPOVCRITLLEVDHGLRVMV
LQVDSGLQOCITVQPKPKPHLPDRIRIVYVTKGFSGTGSENNSTONYKIPPTTK
ALCPILTSRIVTQAPKSTADVSTPDSINILNVDILIRVENIIVILLAGFLSKS
LVFSVLFAVTLRSFVP"

BASE COUNT 258 a 224 c 224 g 199 t

Query Match 1.4%; Score 27; DB 9; Length 905;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1913 ACCGCCAATAAAAAAAAAAAAAA 1939
870 ACCGCCAATAAAAAAAAAAAAAA 896

RESULT 14
LOCUS BC022657 917 bp mRNA linear ROD 16-APR-2003
DEFINITION Mus musculus hypothetical protein LOC225847, mRNA (cDNA clone
IMAGE:4219507), partial cds.
ACCESSION BC022657
VERSION BC022657.1 GI:18490481
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS
MUS musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 917)
Strausberg, R.L., Feingold, E.A., Gronow, L.Y., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buecaw, K.H., Schaefer, C.F., Bhac, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, D., Hsieh, F.,
Diatchenko, L., Martins, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ueding, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitt, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smalinski, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16099-16903 (2002)

JOURNAL
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 917)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcgdpaxli@stanford.edu
Dickson, M., Schmitt, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRK Plate: 43 Row: K Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
FEATURES
source
location/Qualifiers
1..917
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="4219507"
/issue_type="colon, normal, 5 month old male mouse."
/clone_1lb="NCI_CGAP_C024"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..917
/gene="LOC225847"
/db_xref="LOCID:225847"
1..416
/gene="LOC225847"
/codon_start=3
/product="LOC225847 protein"
/protein_id="AAH22657.1"
/db_xref="GI:18490482"
/db_xref="LOCID:225847"
/translation="MRKRLMGILNMLFVSEI,RAATKLTREKYEKEGTLDPKCDY
LKKPSSOKAKMIDIRDEMPKTLACTERPSKSHPOVCRITLLEVDHGLRVMV
LQVDSGLQOCITVQPKPKPHLPDRIRIVYVTKGFSGTGSENNSTONYKIPPTTK
ALCPILTSRIVTQAPKSTADVSTPDSINILNVDILIRVENIIVILLAGFLSKS
LVFSVLFAVTLRSFVP"

BASE COUNT 231 a 255 c 247 g 184 t

Query Match 1.4%; Score 27; DB 10; Length 917;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1913 ACCGCCAATAAAAAAAAAAAAAA 1939
868 ACCGCCAATAAAAAAAAAAAAAA 894

RESULT 15
LOCUS BC027481 963 bp mRNA linear PRI 08-APR-2002
DEFINITION Homo sapiens, clone IMAGE:5107160, mRNA, partial cds.
ACCESSION BC027481
VERSION BC027481.1 GI:20071959
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center

Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 52 Row: f Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES

source

Location/Qualifiers
 1..963
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5107160"
 /issue_type="Cervix, carcinoma"
 /clone_lib="NIH MGC_12"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORTc"
 <1..499

CDS

/codon_start=2
 /product="Unknown (protein for IMAGE:5107160)"
 /protein_id="AAH27481.1"
 /db_xref="GI:20071960"
 /translation="KDGRLMRGLORGASFGSTASALELNRLVQATQATATVYD
 IISPAVPSRSLQDGRGAPADIRGVAKAYPTVREGITDITQICDVASR
 GHQKGLTGAIVGCVIRQLPPTVVKPLITLTAITSSILGGRNQIVPDARKDHALKRS
 DSAQD"

BASE COUNT 190 a 311 c 297 g 165 t
 ORIGIN

Query Match 1.4% Score 27; DB 9; Length 963;
 Best Local Similarity 100.0%; Pred.No. 0.0065;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1913 ACCGCCAAAAAAAAAAAAAAAAAAAA 1939
 Db 920 ACCGCCAAAAAAAAAAAAAAAAAAAA 946

Search completed: January 31, 2004, 06:50:37
 Job time : 7223 secs

ORGANISM: Mus Musculus
US-09-587-436-4

Query Match 1.3%; Score 26; DB 3; Length 1525;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1914 CCGCCAAAAAAAAAAAAAAAAAAAAA 1939
1459 CCGCCAAAAAAAAAAAAAAAAAAAAA 1484

RESULT 3
US-08-927-165A-4
Sequence 4, Application US/08927165A
Patent No. 6410226

GENERAL INFORMATION:

APPLICANT: Kmiec, Eric B.
APPLICANT: Holloman, William K.
APPLICANT: Rice, Michael C.
APPLICANT: Smith, Sheryl T.
APPLICANT: Shu, Zhigang
TITLE OF INVENTION: Mammalian and Human Rec2
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kimeragen, Inc.
STREET: 300 Pleasant Run
CITY: Newtown
STATE: PA
COUNTRY: USA
ZIP: 18940

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PasteSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,165A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hansburg, Daniel
REGISTRATION NUMBER: 36156
REFERENCE/DOCKET NUMBER: 7991-010-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-504-4444
TELEFAX: 215-504-4545

TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1525 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-927-165A-4

Query Match 1.3%; Score 26; DB 4; Length 1525;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1914 CCGCCAAAAAAAAAAAAAAAAAAAAA 1939
1459 CCGCCAAAAAAAAAAAAAAAAAAAAA 1484

RESULT 4
US-09-091-097-26
Sequence 26, Application US/09091097
Patent No. 6432407

GENERAL INFORMATION:

APPLICANT: TAKESAKO, KAZUTOH
APPLICANT: OKADO, TAKASHI
APPLICANT: YAGIHARA, TOMOKO
APPLICANT: KURODA, MASANOBU
APPLICANT: ONISHI, YOSHIMI
APPLICANT: KATO, IKUNOSHIN
APPLICANT: AKIYAMA, KAZUO
APPLICANT: YASUEDA, HIROSHI
TITLE OF INVENTION: ANTIGENIC PROTEIN ORIGINATING IN
TITLE OF INVENTION: MALASSEZIA
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,097
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: WEINER, MARC S.
REGISTRATION NUMBER: 32,181
REFERENCE/DOCKET NUMBER: 1422-0346P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-09-091-097-26

Query Match 1.3%; Score 25; DB 4; Length 263;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1915 CCGCCAAAAAAAAAAAAAAAAAAAAA 1939
200 CCGCCAAAAAAAAAAAAAAAAAAAAA 224

RESULT 5
US-09-325-554-7/c
Sequence 7, Application US/09325554
Patent No. 6410235

GENERAL INFORMATION:
APPLICANT: Weinand, Kurt
APPLICANT: Brand, Joachim
TITLE OF INVENTION: DNA DETECTION BY MEANS OF A STRAND REASSOCIATION COMPLEX
FILE REFERENCE: 024420-00008
CURRENT APPLICATION NUMBER: US/09/325,554
CURRENT FILING DATE: 1999-06-04
PRIOR APPLICATION NUMBER: 198-24-900.4
PRIOR FILING DATE: 1998-06-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent-In version 3.1
SEQ ID NO: 7
LENGTH: 38
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis

FEATURE:
NAME/KEY: misc signal
LOCATION: (1) .. (1)
OTHER INFORMATION: Phosphate linked to biotin via AminoLinker
US-09-325-554-7

Query Match 1.2%; Score 24; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1916 GCCAAGAAAAAAGAAAAA 1939
Db 25 GCCAAGAAAAAAGAAAAA 2

RESULT 6
US-07-820-154A-10/C
Sequence 10, Application US/07820154A
Patent No. 5382425

GENERAL INFORMATION:
APPLICANT: Cochran Ph.D., Mark D
APPLICANT: Junker M.S., David E
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/820.154A
FILING DATE: 19920113
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmid
IMMEDIATE SOURCE:
CLONE: 520-17.5 (Junction B)
FEATURE:
NAME/KEY: CDS
LOCATION: 85..99
OTHER INFORMATION: /function= "translational start of hybrid protein"
OTHER INFORMATION: /product= "N-terminal peptide"
OTHER INFORMATION: /number= 1
OTHER INFORMATION: /standard_name= "Translation of synthetic DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 100..102
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /partial

OTHER INFORMATION: /codon_start= 100
OTHER INFORMATION: /function= "marker enzyme"
OTHER INFORMATION: /product= "Beta-Galactosidase"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "lacZ"
OTHER INFORMATION: /number= 2
OTHER INFORMATION: /citation= ((1))
PUBLICATION INFORMATION:
AUTHORS: Ferrari, Franco A
AUTHORS: Trach, Kathleen
AUTHORS: Hoch, James A
TITLE: Sequence Analysis of the spoB Locus Reveals
TITLE: a Polycistronic Transcription Unit
JOURNAL: J. Bacteriol.
VOLUME: 161
ISSUE: 2
PAGES: 556-562
DATE: Feb.-1985
US-07-820-154A-10

Query Match 1.2%; Score 24; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1916 GCCAAGAAAAAAGAAAAA 1939
Db 78 GCCAAGAAAAAAGAAAAA 55

RESULT 7
US-08-097-554A-10/C
Sequence 10, Application US/08097554A
Patent No. 5869312
GENERAL INFORMATION:
APPLICANT: Cochran Ph.D., Mark D
APPLICANT: Junker M.S., David E
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 112
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097.554A
FILING DATE: July 22, 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmid
IMMEDIATE SOURCE:
CLONE: 520-17.5 (Junction B)

FEATURE:
NAME/KEY: CDS
LOCATION: 85...99 /codon_start= 95
OTHER INFORMATION: /function= "Translational start of hybrid protein"
OTHER INFORMATION: /product= "N-terminal peptide"
OTHER INFORMATION: /number= 1
OTHER INFORMATION: /standard_name= "Translation of synthetic DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 100...102
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /partial
OTHER INFORMATION: /codon_start= 100
OTHER INFORMATION: /function= "marker enzyme"
OTHER INFORMATION: /product= "Beta-galactosidase"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "lacZ"
OTHER INFORMATION: /number= 2
OTHER INFORMATION: /citation= (11)
PUBLICATION INFORMATION:
AUTHORS: Ferrari, Franco A
AUTHORS: Trach, Kathleen
TITLES: Sequence Analysis of the spoB Locus Reveals
TITLE: a Polycistronic Transcription Unit
JOURNAL: J. Bacteriol.
VOLUME: 161
ISSUE: 2
PAGES: 556-562
DATE: Feb.-1985
US-08-097-554A-10

Query Match 1.2%; Score 24; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939
DB 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

RESULT 8
US-08-480-640A-10/c
Sequence 10, Application US/08480640A
Patent No. 6033904
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David B.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,640A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmid
IMMEDIATE SOURCE:
CLONE: 520-17.5 (Junction B)
FEATURE:
NAME/KEY: CDS
LOCATION: 85...99
OTHER INFORMATION: /codon_start= 95
OTHER INFORMATION: /function= "Translational start of hybrid protein"
OTHER INFORMATION: /product= "N-terminal peptide"
OTHER INFORMATION: /number= 1
OTHER INFORMATION: /standard_name= "Translation of synthetic DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 100...102
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /partial
OTHER INFORMATION: /codon_start= 100
OTHER INFORMATION: /function= "marker enzyme"
OTHER INFORMATION: /product= "Beta-galactosidase"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "lacZ"
OTHER INFORMATION: /number= 2
OTHER INFORMATION: /citation= (11)
PUBLICATION INFORMATION:
AUTHORS: Ferrari, Franco A
AUTHORS: Trach, Kathleen
TITLES: Sequence Analysis of the spoB Locus Reveals
TITLE: a Polycistronic Transcription Unit
JOURNAL: J. Bacteriol.
VOLUME: 161
ISSUE: 2
PAGES: 556-562
DATE: Feb.-1985
US-08-480-640A-10

Query Match 1.2%; Score 24; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939
DB 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

RESULT 9
US-08-295-802-10/c
Sequence 10, Application US/08295802
Patent No. 6127163
GENERAL INFORMATION:
APPLICANT: Cochran Ph.D., Mark D
APPLICANT: Junker M.S., David B
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 188
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/295,802
FILING DATE: Herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)977-9550
TELEFAX: (212)664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmid
IMMEDIATE SOURCE:
CLONE: 520-17.5 (Junction B)
FEATURE:
NAME/KEY: CDS
LOCATION: 85..99
OTHER INFORMATION: /codon_start= 85
OTHER INFORMATION: /function= "Translational start of hybrid protein"
OTHER INFORMATION: /product= "N-terminal peptide"
OTHER INFORMATION: /number= 1
OTHER INFORMATION: /standard_name= "Translation of synthetic DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 100..102
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /partial
OTHER INFORMATION: /codon_start= 100
OTHER INFORMATION: /function= "marker enzyme"
OTHER INFORMATION: /product= "Beta-Galactosidase"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "lacZ"
OTHER INFORMATION: /number= 2
OTHER INFORMATION: /citations (11)
PUBLICATION INFORMATION:
AUTHORS: Ferrari, Franco A
AUTHORS: Trach, Kathleen
AUTHORS: Hoch, James A
TITLE: Sequence Analysis of the spoOB Locus Reveals
TITLE: a Polycistronic Transcription Unit
JOURNAL: J. Bacteriol.
VOLUME: 161
ISSUE: 2
PAGES: 556-562
DATE: Feb.-1985
US-08-295-802-10

```

```

Query Match 1.2%; Score 24; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939
DB 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

```

```

RESULT 10
US-08-686-968C-106/c

```

```

Sequence 106, Application US/08686968C
Patent No. 6221351
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David B.
TITLE OF INVENTION: Recombinant Swinepox Virus
FILE REFERENCE: 39119-H/JWL
CURRENT APPLICATION NUMBER: US/08/686,968C
CURRENT FILING DATE: 1996-07-25
NUMBER OF SEQ ID NOS: 231
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 106
LENGTH: 102
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Homology
US-08-686-968C-106

```

```

Query Match 1.2%; Score 24; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939
DB 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

```

```

RESULT 11
US-08-488-237A-10/c
Sequence 10, Application US/08488237A
Patent No. 6251403
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David B.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,237A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmid
IMMEDIATE SOURCE:
CLONE: 520-17.5 (Junction B)

```

FEATURE:
NAME/KEY: CDS
LOCATION: 85..99 /codon_start=85
OTHER INFORMATION: /function="Translational start of hybrid protein"
OTHER INFORMATION: /product="N-terminal peptide"
OTHER INFORMATION: /number=1
OTHER INFORMATION: /standard_name="Translation of synthetic DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 100..102
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /partial
OTHER INFORMATION: /codon_start=100
OTHER INFORMATION: /function="marker enzyme"
OTHER INFORMATION: /product="Beta-galactosidase"
OTHER INFORMATION: /evidence=EXPERIMENTAL
OTHER INFORMATION: /gene="lacZ"
OTHER INFORMATION: /number=2
OTHER INFORMATION: /citation=(11)
PUBLICATION INFORMATION:
AUTHORS: Ferrari, Franco A
AUTHORS: Trach, Kathleen
AUTHORS: Hoch, James A
TITLE: Sequence Analysis of the spoOB locus Reveals
JOURNAL: J. Bacteriol.
VOLUME: 161
ISSUE: 2
PAGES: 556-562
DATE: Feb.-1985
US-08-488-737a-10

Query Match 1.24; Score 24; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939
DB 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

RESULT 12
US-08-375-992a-10/c
Sequence 10, Application US/08375992a
Patent No. 6326975
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 220
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/375,992a
FILING DATE: Herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmid
IMMEDIATE SOURCE:
CLONE: 520-17.5 (Junction B)
FEATURE:
NAME/KEY: CDS
LOCATION: 85..99 /codon_start=85
OTHER INFORMATION: /function="Translational start of hybrid protein"
OTHER INFORMATION: /product="N-terminal peptide"
OTHER INFORMATION: /number=1
OTHER INFORMATION: /standard_name="Translation of synthetic DNA"
FEATURE:
NAME/KEY: CDS
LOCATION: 100..102
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /partial
OTHER INFORMATION: /codon_start=100
OTHER INFORMATION: /function="marker enzyme"
OTHER INFORMATION: /product="Beta-Galactosidase"
OTHER INFORMATION: /evidence=EXPERIMENTAL
OTHER INFORMATION: /gene="lacZ"
OTHER INFORMATION: /number=2
OTHER INFORMATION: /citation=(11)
PUBLICATION INFORMATION:
AUTHORS: Ferrari, Franco A
AUTHORS: Trach, Kathleen
AUTHORS: Hoch, James A
TITLE: Sequence Analysis of the spoOB locus Reveals
JOURNAL: J. Bacteriol.
VOLUME: 161
ISSUE: 2
PAGES: 556-562
DATE: Feb.-1985
US-08-375-992a-10

Query Match 1.24; Score 24; DB 4; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939
DB 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

RESULT 13
US-08-472-679H-10/c
Sequence 10, Application US/08472679H
Patent No. 6497882
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David E.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 267
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pamela G. Salkeld
STREET: 2000 Gallopington Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,679H
FILING DATE: 07-Jun-1995
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Salkeid, Pamela G.
REGISTRATION NUMBER: 38,607
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 298-2135
TELEFAX: (908) 298-5388
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmid
IMMEDIATE SOURCE:
CLONE: 520-17.5 (Junction B)
FEATURE:
NAME/KEY: CDS
LOCATION: 85..99
OTHER INFORMATION: /codon_start= 85/function= "Translational start of hybrid protein"
/product= "N-terminal peptide"/number= 1
/standard_name= "translation of synthetic DNA sequence"
FEATURE:
NAME/KEY: CDS
LOCATION: 100..102
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /partial /codon_start= 100
/function= "marker enzyme"/product= "Beta-Galactosidase"
/evidence= EXPERIMENTAL
/gene= "lacZ"
/number= 2
/citation= ([1])
PUBLICATION INFORMATION:
AUTHORS: Ferrari, Franco A, Trach, Kathleen, Hoch, James A
TITLE: Sequence Analysis of the spoOB Locus Reveals a Polyelectronic Transcription Unit
PATENT NO. 6497882
JOURNAL: J. Bacteriol.
VOLUME: 161
ISSUE: 2
PAGES: 556-562
DATE: Feb.-1985
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-08-472-679H-10
Query Match 1.2%; Score 24; DB 4; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GCCAAAAAAAAAAAAAAAAA 1939
DB 78 GCCAAAAAAAAAAAAAAAAA 55

RESULT 14
PCT-US93-00324-10/C
Sequence 10, Application PC/TUS9300324.
GENERAL INFORMATION:
APPLICANT: Cochran Ph.D., Mark D
APPLICANT: Junker M.S., David E

TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00324
FILING DATE: 19930113
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmid
IMMEDIATE SOURCE:
CLONE: 520-17.5 (Junction B)
FEATURE:
NAME/KEY: CDS
LOCATION: 85..99
OTHER INFORMATION: /codon_start= 85
/function= "Translational start of hybrid protein"
/product= "N-terminal peptide"
/number= 1
/standard_name= "translation of synthetic DNA sequence"
FEATURE:
NAME/KEY: CDS
LOCATION: 100..102
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /partial
OTHER INFORMATION: /codon_start= 100
/function= "marker enzyme"
/product= "Beta-Galactosidase"
/evidence= EXPERIMENTAL
/gene= "lacZ"
/number= 2
/citation= ([1])
PUBLICATION INFORMATION:
AUTHORS: Ferrari, Franco A
AUTHORS: Trach, Kathleen
AUTHORS: Hoch, James A
TITLE: Sequence Analysis of the spoOB Locus Reveals a Polyelectronic Transcription Unit
JOURNAL: J. Bacteriol.
VOLUME: 161
ISSUE: 2
PAGES: 556-562
DATE: Feb.-1985
PCT-US93-00324-10
Query Match 1.2%; Score 24; DB 5; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.47;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939

Db 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

RESULT 15

US-07-820-154A-32/c

Sequence 32, Application US/07820154A

Patent No. 5382425

GENERAL INFORMATION:

APPLICANT: Cochran Ph.D., Mark D.

APPLICANT: Junker M.S., David E.

TITLE OF INVENTION: Recombinant Swinepox Virus

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESSEE: John P. White

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/820,154A

FILING DATE: 19920113

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: White, John P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 977-9550

TELEFAX: (212) 664-0525

TELEX: 422523

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 108 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Plasmid

IMMEDIATE SOURCE:

CLONE: 538-46.26 (Junction B)

FEATURE:

NAME/KEY: exon

LOCATION: 88..102

OTHER INFORMATION: /codon_start= 88

OTHER INFORMATION: /function= "Translational start of hybrid protein"

OTHER INFORMATION: /product= "N-terminal peptide"

OTHER INFORMATION: /number= 1

OTHER INFORMATION: /standard_name= "Translation of synthetic DNA sequence"

FEATURE:

NAME/KEY: CDS

LOCATION: 103..108

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /partial

OTHER INFORMATION: /codon_start= 103

OTHER INFORMATION: /product= "NDV Hemagglutinin-Neuraminidase"

OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /gene= "HN"

OTHER INFORMATION: /number= 2

US-07-820-154A-32

Query Match

1.2% Score 24; DB 1; Length 108;

Best Local Similarity 100.0%; Pred. No. 0.47;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939

Db 81 GCCAAAAAAAAAAAAAAAAAAAAA 58

Search completed: January 31, 2004, 08:03:22

Job time : 128 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2004, 04:06:41, Search time 550 Seconds

(without alignments)
9516,746 Million cell updates/sec

Title: US-10-017-085A-205

Perfect score: 1939
Sequence: 1 cgcctcgcgccttcg9ag9ctc.....aaaaaaaaaaaaaaaa 1939

Scoring table: OLIGO NUC

Gapop 60.0, Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size: 10

Total number of hits satisfying chosen parameters: 1552787

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1939	100.0	1939	20	AAZ34065 Human PRO853 nucle
2	1939	100.0	1939	20	AAZ34065 Human PRO853 nucle
3	1939	100.0	1939	21	AAZ34065 Human PRO853 nucle
4	1939	100.0	1939	21	AAZ34065 Human PRO853 nucle
5	1939	100.0	1939	25	ABX92437 Human PRO853 nucle
6	1803	93.0	1887	24	ABL90069 Human polynucleoti
7	1497	77.2	1779	22	ABD09338 Human drug metabol
8	779	40.2	1098	25	ABZ68100 Human secreted pro

9	698	36.0	1140	22	AAZ34065	DNA encoding human
10	698	36.0	1140	25	AAZ34065	Human secreted pro
11	396	20.4	598	21	AAZ34065	Human ORFX ORF1952
12	200	10.3	397	22	AAZ34065	Novel human polyu
13	92	4.7	186	16	AAZ34065	Human gene signatu
14	88	4.5	175	24	AAZ34065	Human benign prost
15	51	2.6	51	22	AAZ34065	Human silent SNP c
16	45	2.3	45	20	AAZ34065	Human PRO853 hybr
17	45	2.3	45	21	AAZ34065	Human PRO853 hybr
18	45	2.3	45	21	AAZ34065	Probe used to scre
19	45	2.3	45	25	AAZ34065	Human PRO DNA prob
20	27	1.4	1481	21	AAZ34065	Human secreted pro
21	27	1.4	2093	21	AAZ34065	Human ORFX ORF3039
22	26	1.3	159	23	AAZ34065	Human prostate exp
23	26	1.3	159	23	AAZ34065	Human prostate exp
24	26	1.3	179	23	AAZ34065	Human prostate exp
25	26	1.3	207	24	AAZ34065	Human prostate exp
26	26	1.3	282	22	AAZ34065	Lung cancer relate
27	26	1.3	295	23	AAZ34065	Human breast cance
28	26	1.3	356	22	AAZ34065	Human prostate exp
29	26	1.3	359	22	AAZ34065	Human cervical can
30	26	1.3	360	22	AAZ34065	Human cervical can
31	26	1.3	407	22	AAZ34065	Human polynucleoti
32	26	1.3	419	23	AAZ34065	Human prostate exp
33	26	1.3	481	22	AAZ34065	Human breast cance
34	26	1.3	495	22	AAZ34065	Human breast cance
35	26	1.3	495	22	AAZ34065	Human foetal liver
36	26	1.3	495	22	AAZ34065	Human brain expres
37	26	1.3	495	22	AAZ34065	Human bone marrow
38	26	1.3	495	22	AAZ34065	Probe #152 for gen
39	26	1.3	495	22	AAZ34065	Probe #154 used to
40	26	1.3	495	22	AAZ34065	Probe #149 used to
41	26	1.3	495	22	AAZ34065	Human liver single
42	26	1.3	495	24	AAZ34065	Human genome deriv
43	26	1.3	570	24	AAZ34065	Oligonucleotide fo
44	26	1.3	570	24	AAZ34065	Oligonucleotide fo
45	26	1.3	831	20	AAZ34065	Human secreted pro

ALIGNMENTS

RESULT 1	
AAZ34065	AAZ34065 standard; CDNA; 1939 BP.
ID	AAZ34065
AC	AAZ34065
XX	07-DEC-1999 (first entry)
DT	Human PRO853 nucleotide sequence.
XX	
DE	Human; PRO: EST: expressed sequence tag; PCR primer; hybridization;
XX	probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW	secreted protein; transmembrane protein; ss.
KM	
XX	
OS	Homo sapiens.
XX	
PN	W09946281-A2.
PD	16-SEP-1999.
XX	
PF	08-MAR-1999; 99WC-US05028.
XX	
PR	10-MAR-1998; 98US-0077450.
PR	11-MAR-1998; 98US-0077632.
PR	11-MAR-1998; 98US-0077641.
PR	11-MAR-1998; 98US-0077649.
PR	12-MAR-1998; 98US-0077791.
PR	13-MAR-1998; 98US-0078004.
PR	17-MAR-1998; 98US-0040220.
PR	20-MAR-1998; 98US-0078886.
PR	20-MAR-1998; 98US-0078910.

PR 20-MAR-1998; 98US-0078936.
 PR 20-MAR-1998; 98US-0078939.
 PR 25-MAR-1998; 98US-0079294.
 PR 25-MAR-1998; 98US-0079656.
 PR 27-MAR-1998; 98US-0079663.
 PR 27-MAR-1998; 98US-0079664.
 PR 27-MAR-1998; 98US-0079689.
 PR 27-MAR-1998; 98US-0079728.
 PR 27-MAR-1998; 98US-0079786.
 PR 30-MAR-1998; 98US-0079920.
 PR 30-MAR-1998; 98US-0079923.
 PR 31-MAR-1998; 98US-0080105.
 PR 31-MAR-1998; 98US-0080107.
 PR 31-MAR-1998; 98US-0080165.
 PR 31-MAR-1998; 98US-0080194.
 PR 01-APR-1998; 98US-0080327.
 PR 01-APR-1998; 98US-0080328.
 PR 01-APR-1998; 98US-0080333.
 PR 01-APR-1998; 98US-0080334.
 PR 08-APR-1998; 98US-0081049.
 PR 08-APR-1998; 98US-0081070.
 PR 08-APR-1998; 98US-0081071.
 PR 09-APR-1998; 98US-0081195.
 PR 09-APR-1998; 98US-0081203.
 PR 09-APR-1998; 98US-0081229.
 PR 15-APR-1998; 98US-0081817.
 PR 15-APR-1998; 98US-0081838.
 PR 15-APR-1998; 98US-0081952.
 PR 15-APR-1998; 98US-0081955.
 PR 21-APR-1998; 98US-0082568.
 PR 21-APR-1998; 98US-0082569.
 PR 22-APR-1998; 98US-0082700.
 PR 22-APR-1998; 98US-0082704.
 PR 23-APR-1998; 98US-0082804.
 PR 23-APR-1998; 98US-0082767.
 PR 27-APR-1998; 98US-0082796.
 PR 28-APR-1998; 98US-0083336.
 PR 28-APR-1998; 98US-0083342.
 PR 29-APR-1998; 98US-0083392.
 PR 29-APR-1998; 98US-0083495.
 PR 29-APR-1998; 98US-0083496.
 PR 29-APR-1998; 98US-0083499.
 PR 29-APR-1998; 98US-0083500.
 PR 29-APR-1998; 98US-0083545.
 PR 29-APR-1998; 98US-0083554.
 PR 29-APR-1998; 98US-0083558.
 PR 30-APR-1998; 98US-0083742.
 PR 05-MAY-1998; 98US-0084356.
 PR 06-MAY-1998; 98US-0084414.
 PR 06-MAY-1998; 98US-0084441.
 PR 07-MAY-1998; 98US-0084598.
 PR 07-MAY-1998; 98US-0084600.
 PR 07-MAY-1998; 98US-0084627.
 PR 07-MAY-1998; 98US-0084637.
 PR 07-MAY-1998; 98US-0084639.
 PR 07-MAY-1998; 98US-0084640.
 PR 07-MAY-1998; 98US-0084643.
 PR 13-MAY-1998; 98US-0085233.
 PR 13-MAY-1998; 98US-0085338.
 PR 13-MAY-1998; 98US-0085339.
 PR 13-MAY-1998; 98US-0085373.
 PR 15-MAY-1998; 98US-0085579.
 PR 15-MAY-1998; 98US-0085580.
 PR 15-MAY-1998; 98US-0085582.
 PR 15-MAY-1998; 98US-0085689.
 PR 15-MAY-1998; 98US-0085697.
 PR 15-MAY-1998; 98US-0085700.
 PR 15-MAY-1998; 98US-0085704.
 PR 18-MAY-1998; 98US-0086023.
 PR 22-MAY-1998; 98US-0086392.
 PR 22-MAY-1998; 98US-0086414.
 PR 22-MAY-1998; 98US-0086430.

PR 22-MAY-1998; 98US-0086486.
 PR 28-MAY-1998; 98US-0087098.
 PR 28-MAY-1998; 98US-0087106.
 PR 28-MAY-1998; 98US-0087208.
 PR 30-JUL-1998; 98US-0094651.
 PR 11-SEP-1998; 98US-0100038.
 XX (GETH) GENENTECH INC.
 XX PA
 XX Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J,
 PI WPI, 1999-551358/46.
 XX P-PSDB; AA41715.
 DR
 XX
 XX New secreted and transmembrane polypeptides and their polynucleotides,
 PT useful for treating blood coagulation disorders, cancers and cellular
 PT adhesion disorders -
 XX
 XX Claim 2, Fig 74, 530pp, English.
 PS
 XX The present invention describes secreted and transmembrane polypeptides
 CC and their polynucleotides. The nucleotide sequences are useful as
 CC sources of probes, primers, for chromosome mapping, and for generation
 CC of antisense sequences. They can also be used to create transgenic
 CC animals. The proteins can be used to treat a variety of diseases and
 CC disorders, depending on their function. Diseases that may be treated
 CC include blood coagulation disorders, cancers and cellular adhesion
 CC disorders. They may also be used to raise antibodies. AA233891 to
 CC AA234338, and AA41685 to AA41774 represent polynucleotide and
 CC polypeptide sequence given in the exemplification of the present
 CC invention.
 XX
 XX Sequence 1939 BP; 367 A; 549 C; 624 G; 399 T; 0 other;
 SQ
 Query Match 100.0%; Score 1939; DB 20; Length 1939;
 Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;
 Matches 1939; Conservative 0; Mismatches 0
 QY 1 CGGCTCCGCTTCGAGGCTGACGCGCCGCGGCGCTTCAGAGCTGTGACGAGCGGAT 60
 DB 1 CGGCTCCGCTTCGAGGCTGACGCGCCGCGGCGCTTCAGAGCTGTGACGAGCGGAT 60
 QY 61 CGGACGCCGCTGCGCGGCATCCAGGGCGGTGCGGGCCTTGCGGAGCGGAGCGGC 120
 DB 61 CGGACGCCGCTGCGCGGCATCCAGGGCGGTGCGGGCCTTGCGGAGCGGAGCGGC 120
 QY 121 GGGCGGATGAGAGGCGCTGCGGGCGCGGGGCTTGCTGCGGGGCTTACGTGCTGT 180
 DB 121 GGGCGGATGAGAGGCGCTGCGGGCGCGGGGCTTGCTGCGGGGCTTACGTGCTGT 180
 QY 181 CTACTACAACCTGTGTAAGCCCGCGTGTGCGGCGGATGAGGCACTTGCGGGCGGAC 240
 DB 181 CTACTACAACCTGTGTAAGCCCGCGTGTGCGGCGGATGAGGCACTTGCGGGCGGAC 240
 QY 241 GGGCGGTGTCAGGGCGGCCAAGCGCGCATCGGAAGATGACGGCGCTGAGCTGGCGG 300
 DB 241 GGGCGGTGTCAGGGCGGCCAAGCGCGCATCGGAAGATGACGGCGCTGAGCTGGCGG 300
 QY 301 GGGCGGAGGCGGCTGTGCTGCGGCTGCGGCGGAGCGGCGGAGCGGCGGCTGCTT 360
 DB 301 GGGCGGAGGCGGCTGTGCTGCGGCTGCGGCGGAGCGGCGGAGCGGCGGCTGCTT 360
 QY 361 CGACCTCCGCGAGAGAGTGGGAACATGAGGTCACTTCATGAGCTTGAGACTTGGCAG 420
 DB 361 CGACCTCCGCGAGAGAGTGGGAACATGAGGTCACTTCATGAGCTTGAGACTTGGCAG 420
 QY 421 TCTGGCTGTGCTGCGGCGCTTGGCACTGCTTCTGAGCTTGAGCGCGGTTGACAT 480
 DB 421 TCTGGCTGTGCTGCGGCGCTTGGCACTGCTTCTGAGCTTGAGCGCGGTTGACAT 480
 QY 481 CCTCATCCCAATGCGCGGTATGAGTCTGTGCGGAGCCCGGAGCGGTTAACCTGCT 540
 DB 481 CCTCATCCCAATGCGCGGTATGAGTCTGTGCGGAGCCCGGAGCGGTTAACCTGCT 540

QY	541	GCTTCGGGATGAACCAATATCGGATCCCTTCTCTCTGACATATGCGCTGACCTTGCTGATCA	600
Db	541	GCTTCGGGATGAACCAATATCGGATCCCTTCTCTCTGACATATGCGCTGACCTTGCTGATCA	600
QY	601	GGCATATGCCCTTAAGCCCGCGTGGTGGTGAAGCTTCAAGCTGACCTGACCTGATCGGAGACGCTT	660
Db	601	GGCATATGCCCTTAAGCCCGCGTGGTGGTGAAGCTTCAAGCTGACCTGACCTGATCGGAGACGCTT	660
QY	661	TGACTTTAAAGGCGCTGAGACCGCCGACATGATGGGCTGGGGGAGAGAGCTGCGGGGCAATATGC	720
Db	661	TGACTTTAAAGGCGCTGAGACCGCCGACATGATGGGCTGGGGGAGAGAGCTGCGGGGCAATATGC	720
QY	721	TGACACTAAGCTGAGCTAATATGATCTGTTTGGCCCGGAGAGCTGACCAACAGCTTGAAGGCGAC	780
Db	721	TGACACTAAGCTGAGCTAATATGATCTGTTTGGCCCGGAGAGCTGACCAACAGCTTGAAGGCGAC	780
QY	781	TGGCGTCAACCTGCTATATGACAGCCCAACCCAGGGGCTGTGTAATCTGGAGAGCTGTTCTGCGCCA	840
Db	781	TGGCGTCAACCTGCTATATGACAGCCCAACCCAGGGGCTGTGTAATCTGGAGAGCTGTTCTGCGCCA	840
QY	841	TGTTCTGATGAGCTGCGGCCCACTTTTGGCCCATTTGCGCCCATTTGGCTTGGCTGGTCTCGGGACCC	900
Db	841	TGTTCTGATGAGCTGCGGCCCACTTTTGGCCCATTTGCGCCCATTTGGCTTGGCTGGTCTCGGGACCC	900
QY	901	AAGAGGGGGGTGCGCAGACACCCCTGTATTTGTGCTTACAAAGAGGGCATGAGCCCTCTAG	960
Db	901	AAGAGGGGGGTGCGCAGACACCCCTGTATTTGTGCTTACAAAGAGGGCATGAGCCCTCTAG	960
QY	961	TGGAGATATTTTGGCAATGTCGCAATGATGAGGAGAGGTGCTCAGCTGCGCCGAGACGACCG	1020
Db	961	TGGAGATATTTTGGCAATGTCGCAATGATGAGGAGAGGTGCTCAGCTGCGCCGAGACGACCG	1020
QY	1021	GGCAGCCCATCGGCTATGAGGAGGCCAGCAAGAGGCTGGAGAGGCTTGGGCTTGGGAGAGA	1080
Db	1021	GGCAGCCCATCGGCTATGAGGAGGCCAGCAAGAGGCTGGAGAGGCTTGGGCTTGGGAGAGA	1080
QY	1081	TGCTGAAACCCGATGAAGAAGCCCGCACTTGAAGAGCTCAAGAGGCCCATTTTCTTAAAGAC	1140
Db	1081	TGCTGAAACCCGATGAAGAAGCCCGCACTTGAAGAGCTCAAGAGGCCCATTTTCTTAAAGAC	1140
QY	1141	CCCCCAACCTGAGAGAGCCCAAGTGTCTCAACCTTACACCTTACCCAGCCCTGAGAGCTCAACAGA	1200
Db	1141	CCCCCAACCTGAGAGAGCCCAAGTGTCTCAACCTTACACCTTACCCAGCCCTGAGAGCTCAACAGA	1200
QY	1201	TTTGTCTAAGATGACGACCCGAATTCAGGCTTAAAGTTAGGCTGATGATCAGCTCTCCTA	1266
Db	1201	TTTGTCTAAGATGACGACCCGAATTCAGGCTTAAAGTTAGGCTGATGATCAGCTCTCCTA	1266
QY	1261	AACCTCAGGCGAGAGATGCTTGCATGAGCACTTCAATGATCTTGAAGAACTTGGGATATGCTG	1320
Db	1261	AACCTCAGGCGAGAGATGCTTGCATGAGCACTTCAATGATCTTGAAGAACTTGGGATATGCTG	1320
QY	1321	TGAGGCGCATGCCCTGAGACATGACGGGTTTGTGATCTTGAACCTCCGATGTTACTTCTGCG	1380
Db	1321	TGAGGCGCATGCCCTGAGACATGACGGGTTTGTGATCTTGAACCTCCGATGTTACTTCTGCG	1380
QY	1381	GGCCCCAAGCTGTGCTCCCTGAGACATCTCTTCTCTGTTGAAGAAATATGCGGTATATTT	1440
Db	1381	GGCCCCAAGCTGTGCTCCCTGAGACATCTCTTCTCTGTTGAAGAAATATGCGGTATATTT	1440
QY	1441	TCTTCTGAGAGTGCATGTAACCCGACATGAGAGATATGAGGGTATGCTAGACATCTGTGCT	1500
Db	1441	TCTTCTGAGAGTGCATGTAACCCGACATGAGAGATATGAGGGTATGCTAGACATCTGTGCT	1500
QY	1501	TCTGGAAATTTGAGATGATATTTTCAAGGCCCAACCTTATTTGATCTGATCAAGCTCTG	1566
Db	1501	TCTGGAAATTTGAGATGATATTTTCAAGGCCCAACCTTATTTGATCTGATCAAGCTCTG	1566
QY	1561	GAGCAGAGCAGGAGGTTTGCATATGTATGACTGCCAACATTTGAGATTTAGTGAACCTGA	1620
Db	1561	GAGCAGAGCAGGAGGTTTGCATATGTATGACTGCCAACATTTGAGATTTAGTGAACCTGA	1620

QY	1621	TCCCTTTGCAACCGCTTAGCTAAGGTAGTAAATTACCCCATGTATTATGAAGCGAATTA	1680
Db	1621	TCCCTTTGCAACCGCTTAGCTAAGGTAGTAAATTACCCCATGTATTATGAAGCGAATTA	1680
QY	1681	GCGCTCCGAGCTAAGGACCTCGCTTAGAGTCTCAACAGTGAATGAGAGAGGCGCTTGGAT	1740
Db	1681	GCGCTCCGAGCTAAGGACCTCGCTTAGAGTCTCAACAGTGAATGAGAGAGGCGCTTGGAT	1740
QY	1741	CTGAACCCMAAGGTTCTGAGCGCAGAGGCGGACCTGCGTAAAGTGGTCTTAGAAGTAGT	1800
Db	1741	CTGAACCCMAAGGTTCTGAGCGCAGAGGCGGACCTGCGTAAAGTGGTCTTAGAAGTAGT	1800
QY	1801	CAGGCGAAGGACACTGTGTATGTAGAGTGCCCCATATGGAGTAAAGGGAGCGCCCTTCGGGCGG	1860
Db	1801	CAGGCGAAGGACACTGTGTATGTAGAGTGCCCCATATGGAGTAAAGGGAGCGCCCTTCGGGCGG	1860
QY	1861	ATGCAGGCGTGGGCTCATCTGTATCTGAAGCCCTCGGATTAAGCGCGTTGACCGCCA	1920
Db	1861	ATGCAGGCGTGGGCTCATCTGTATCTGAAGCCCTCGGATTAAGCGCGTTGACCGCCA	1920
QY	1921	AAAAAAAAAAAAAAAAAAAA	1939
Db	1921	AAAAAAAAAAAAAAAAAAAA	1939

XX	RESULT 2
XX	.AAx87266
ID	AAx87266 standard; cDNA, 1939 BP.
AC	AAx87266;
DT	27-SEP-1999 (first entry)
DE	cDNA clone encoding human PRO853, amplified in tumour cells.
KW	PRO853; UNQ419; tumour; cancer; diagnosis; therapy; human; ds..
OS	Homo sapiens.
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	128..1261
FT	/*tag= a'
FT	sig_peptide
FT	128..175
FT	/*tag= b
FT	mat_peptide
FT	176..1258
FT	/*tag= c
PN	WO935170-A2.
PD	15-JUL-1999.
PF	05-JAN-1999; 99NO-US00106.
PR	20-NOV-1998; 98US-0109304.
PR	05-JAN-1998; 98US-0070440.
PR	29-APR-1998; 98US-0083500.
PR	22-MAY-1998; 98US-0086614.
PR	10-JUN-1998; 98US-0088742.
PR	10-NOV-1998; 98US-0107783.
PA	(GETH) GENENTECH INC.
PI	Bosestin D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA,
PI	Roy MA, Wood WT;
DR	WPI; 1999-430385/36.
DR	P-PDSB; AAy06489.
PT	Antibody against proteins expressed in neoplastic cells, useful for
PS	tumor diagnosis and treatment
XX	Example 1; Fig 25A-B; 16zpp; English.

CC This is the nucleotide sequence of cDNA clone DNA4827 (ATCC 209812)
CC coding for human PRO83 (ONQ419) (see AY06489). The clone was
CC isolated from a human foetal kidney tissue library. Amplification
CC of DNA4827 (chromosome 17) occurs in various tumours, especially,
CC colon tumours, suggesting an association with tumour formation or
CC growth. Antagonists (e.g. antibodies) directed to PRO83 may have
CC use in cancer therapy. The invention identifies 14 genes (see
CC AX87254-67) that are amplified in the genome of tumour cells. Such
CC amplification is expected to be associated with overexpression of
CC the gene product and to contribute to tumorigenesis. The encoded
CC proteins (see AY06477-90) may be useful targets for the diagnosis
CC and/or treatment (including prevention) of certain cancers, and may
CC act as predictors of the prognosis of tumour treatment.

XX Sequence 1939 BP; 367 A; 549 C; 624 G; 399 T; 0 other;

Query Match 100.0%; Score 1939; DB 20; Length 1939;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTCCGCTTGGAGAGCTGACGCGCCGCGCGCTTCCAGGCTGTGCAAGGCGGAT 60
DB 1 CGCTCCGCTTGGAGAGCTGACGCGCCGCGCGCTTCCAGGCTGTGCAAGGCGGAT 60
QY 61 CGGAGCGCGCTGGCGGCGATCCAGGCGGCTGGCGGCGCGCGGAGCGCGGAGCGC 120
DB 61 CGGAGCGCGCTGGCGGCGATCCAGGCGGCTGGCGGCGCGCGGAGCGCGGAGCGC 120
QY 121 GCGCGGATGAGAGCGCTGTGCTGTGGCGCGGCGGCTGTGCTGTGGCGCTTACGCTTGT 180
DB 121 GCGCGGATGAGAGCGCTGTGCTGTGGCGCGGCGGCTGTGCTGTGGCGCTTACGCTTGT 180
QY 181 CTACTAACCTGTGTGAAGGCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCG 240
DB 181 CTACTAACCTGTGTGAAGGCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCG 240
QY 241 GCGCGGATGAGAGCGCTGTGCTGTGGCGCGGCGGCTGTGCTGTGGCGCTTACGCTTGT 300
DB 241 GCGCGGATGAGAGCGCTGTGCTGTGGCGCGGCGGCTGTGCTGTGGCGCTTACGCTTGT 300
QY 301 CCGGAGAGCGCGCTGTGCTGTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 360
DB 301 CCGGAGAGCGCGCTGTGCTGTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 360
QY 361 CGACTCCGCGGAGAGAGTGGGAAACAATGAGGTCATCTTATGCGCTTGAAGCTTGGCAG 420
DB 361 CGACTCCGCGGAGAGAGTGGGAAACAATGAGGTCATCTTATGCGCTTGAAGCTTGGCAG 420
QY 421 TCGGCGCTGGTGGCGGCGCTTGGCACTGCTTCTGAGCTGTGAGGCAAGGTGAGCAT 480
DB 421 TCGGCGCTGGTGGCGGCGCTTGGCACTGCTTCTGAGCTGTGAGGCAAGGTGAGCAT 480
QY 481 CCTCATCAACAATCCGCGTATCAAGTCTGTGCGCGGAGCGGAGCGGCTTGAAGCTTGT 540
DB 481 CCTCATCAACAATCCGCGTATCAAGTCTGTGCGCGGAGCGGAGCGGCTTGAAGCTTGT 540
QY 541 GCTTCGCGTGAACAATATCGGTCCCTTCTGCTGACACATCTGTGCTGCTTGGCTGA 600
DB 541 GCTTCGCGTGAACAATATCGGTCCCTTCTGCTGACACATCTGTGCTGCTTGGCTGA 600
QY 601 GCGATGTGCGCGCTGAGCGCGGCTGTGCTGTGAGCTTCAAGCTGAGCGGAGCGTCT 660
DB 601 GCGATGTGCGCGCTGAGCGCGGCTGTGCTGTGAGCTTCAAGCTGAGCGGAGCGTCT 660
QY 661 TGAATTGAGCGCTGAGCGCGGCTGTGCTGTGAGCTTCAAGCTGAGCGGAGCGTCT 720
DB 661 TGAATTGAGCGCTGAGCGCGGCTGTGCTGTGAGCTTCAAGCTGAGCGGAGCGTCT 720
QY 721 TGAACAATGAGCTGAGCTTATGCTTGTGCGGAGCGCTGAGCGGAGCGTCTGAGC 780
DB 721 TGAACAATGAGCTGAGCTTATGCTTGTGCGGAGCGCTGAGCGGAGCGTCTGAGC 780
QY 781 TGGGATCACTGCTATGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGC 840

DB 781 TGGGATCACTGCTATGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGC 840
QY 841 TGTCTCTGATGAGCTGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGC 900
DB 841 TGTCTCTGATGAGCTGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGC 900
QY 901 AAGAGGAGGAGTGGCAGACACCCCTGATATGCTTCAAGAGGAGCGGAGCGGAGC 960
DB 901 AAGAGGAGGAGTGGCAGACACCCCTGATATGCTTCAAGAGGAGCGGAGCGGAGC 960
QY 961 TGGGATCACTGCTATGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGC 1020
DB 961 TGGGATCACTGCTATGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGC 1020
QY 1021 GCGAGCGGAGTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGC 1080
DB 1021 GCGAGCGGAGTGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGC 1080
QY 1081 TGTGAACTCGATGAAAGACCCCGAGCTGAGAGCTGAGAGCGGAGCGGAGC 1140
DB 1081 TGTGAACTCGATGAAAGACCCCGAGCTGAGAGCTGAGAGCGGAGCGGAGC 1140
QY 1141 CCGCGACCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGC 1200
DB 1141 CCGCGACCTGAGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGC 1200
QY 1201 TTTGTCTAAGATAGCGACCGAATTCAGGCTTAAAGTGAAGCTGAGAGCTGCTCTA 1260
DB 1201 TTTGTCTAAGATAGCGACCGAATTCAGGCTTAAAGTGAAGCTGAGAGCTGCTCTA 1260
QY 1261 ACCCTGAGCGGAGAGTCTTGCATGAGCACTTATGCTTGAAGCTTGAAGCTGCT 1320
DB 1261 ACCCTGAGCGGAGAGTCTTGCATGAGCACTTATGCTTGAAGCTTGAAGCTGCT 1320
QY 1321 TGAAGGAGTGGCTTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1380
DB 1321 TGAAGGAGTGGCTTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1380
QY 1381 GCGCCGAGAGTGGCTTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1440
DB 1381 GCGCCGAGAGTGGCTTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1440
QY 1441 TCTTCTGAGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1500
DB 1441 TCTTCTGAGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1500
QY 1501 TCTGGAATTTGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1560
DB 1501 TCTGGAATTTGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1560
QY 1561 GAGCAGAGGAGGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1620
DB 1561 GAGCAGAGGAGGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1620
QY 1621 TCCCTTGAAGCGGCTGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTT 1680
DB 1621 TCCCTTGAAGCGGCTGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTT 1680
QY 1681 GGTCTCCGAGTGAAGGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTT 1740
DB 1681 GGTCTCCGAGTGAAGGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTT 1740
QY 1741 CTGAACCAAGGCTGAGGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTT 1800
DB 1741 CTGAACCAAGGCTGAGGAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTT 1800
QY 1801 CAGGAGAGGAGTGAAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1860
DB 1801 CAGGAGAGGAGTGAAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1860
QY 1861 ATGAGAGGAGTGAAGTGAAGTGAAGCACTGAGCGGCTTGAAGCTTGAAGCTTGAAGCTT 1920

DB 1861 ATGACGAGGCTGAGGCTATCTGATCTGAAGCCCTCTGAAATTAAGCGCTTGACCGCCAA 1920
 QY 1921 AAAAAAAAAAAAAAAAAA 1939
 XX |||||
 DB 1921 AAAAAAAAAAAAAAAAAA 1939
 RESULT 3
 AACT8501
 ID AACT8501 standard; cDNA, 1939 BP.
 XX
 AC AACT8501;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human PRO853 (UNQ419) nucleotide sequence SEQ ID NO:205.
 XX
 KW Human, secreted protein; transmembrane protein; PRO; EST; cytosolic;
 XX expressed sequence tag; detection; cancer; ss.
 OS Homo sapiens.
 XX
 PN WO200053756-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 18-FEB-2000; 2000WO-US04341.
 XX
 PR 08-MAR-1999; 99WO-US05028.
 PR 12-MAR-1999; 99US-0123857.
 PR 29-MAR-1999; 99US-0126773.
 PR 21-APR-1999; 99US-0130232.
 PR 28-APR-1999; 99US-0131445.
 PR 14-MAY-1999; 99US-0134287.
 PR 23-JUN-1999; 99US-0141037.
 PR 26-JUL-1999; 99US-0145698.
 PR 29-OCT-1999; 99US-0162506.
 PR 30-NOV-1999; 99WO-US28513.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28565.
 PR 16-DEC-1999; 99WO-US30095.
 PR 30-DEC-1999; 99WO-US31243.
 PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 06-JAN-2000; 2000WO-US00376.
 XX
 XX (GENTH) GENTECH INC.
 PA
 PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,
 PI Ferreira N, Flivaroff E, Fong S, Gao W, Gerber H, Gerritsen ME,
 PI Goddard AJ, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ,
 PI Klijavin IJ, Kuo SS, Napier MA, Pan J, Paoni NP, Roy MA,
 PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood MT;
 XX
 DR WPI, 2000-611443/58.
 XX
 DR P-PSDB; AAB44271.
 XX
 PT Novel PRO polypeptides and polynucleotides used in detection methods,
 PT to target bioactive molecules to specific cells, and to modulate
 PT cellular activities -
 XX
 PS Claim 2, Fig 74; 636pp; English.
 XX
 CC AACT8458 to AACT8559 represent polynucleotide and EST (expressed
 CC sequence tag) sequences which encode secreted or transmembrane PRO
 CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic
 CC activity. The polynucleotides and polypeptides can be used for detecting
 CC the presence of PRO polypeptides in samples, for linking bioactive
 CC molecules to cells and for modulating biological activities of cells,
 CC using the polypeptides for specific targeting. The polypeptide targeting
 CC can be used to kill the target cells, e.g. for the treatment of cancers.
 CC The polypeptide pairs provide specific targeting of bioactive molecules

CC to cells. AACT8600 to AACT8987 represent PCR primers and probes used in
 CC the isolation of the PRO polynucleotide sequences.
 SQ Sequence 1939 BP; 367 A; 549 C; 624 G; 399 T; 0 other;
 Query Match 100.0%; Score 1939; DB 21; Length 1939;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCCTCCGCTTCGAGAGGCTGAGCGCCCGGCGCGCTTCAGAGGCTGTGACAGGGGAGT 60
 DB 1 CGCCTCCGCTTCGAGAGGCTGAGCGCCCGGCGCGCTTCAGAGGCTGTGACAGGGGAGT 60
 QY 61 CGGACCGCGCTGCGCGGCGATCAGAGCGGTCGCGGCGCTGCGCGGAGCCGAGAGCGC 120
 DB 61 CGGACCGCGCTGCGCGGCGATCAGAGCGGTCGCGGCGCTGCGCGGAGCCGAGAGCGC 120
 QY 121 GCGCGGATGAGAGCGCTGCTGCTGCGCGCGGCGGCTTCTGCTGCGCGCTTACGCTTGT 180
 DB 121 GCGCGGATGAGAGCGCTGCTGCTGCGCGCGGCGGCTTCTGCTGCGCGCTTACGCTTGT 180
 QY 181 CTACTACAACTGAGAGGC 240
 DB 181 CTACTACAACTGAGAGGC 240
 QY 241 GCGCGTGTGTCAGCG 300
 DB 241 GCGCGTGTGTCAGCG 300
 QY 301 CGCGGAGAGCG 360
 DB 301 CGCGGAGAGCG 360
 QY 361 CGACCTCGCGCGAGAGAGGAGAGCAATAGAGTCACTTATGAGCGCTTGAAGCTTGGCAG 420
 DB 361 CGACCTCGCGCGAGAGAGGAGAGCAATAGAGTCACTTATGAGCGCTTGAAGCTTGGCAG 420
 QY 421 TCTGCGCTCGAGCGCGCGCTTTCGCACTGCTCTTTCGAGCTGAGCCAGCGTGGACAT 480
 DB 421 TCTGCGCTCGAGCGCGCGCTTTCGCACTGCTCTTTCGAGCTGAGCCAGCGTGGACAT 480
 QY 481 CCTATCCAAATGCGCGGTATGATTCGTGTGCGCGGAGCCGTAAGCGCTTAACTGCT 540
 DB 481 CCTATCCAAATGCGCGGTATGATTCGTGTGCGCGGAGCCGTAAGCGCTTAACTGCT 540
 QY 541 GCTTGGGAGAACATATGCGGCTTCTGCTGCAACATGCTGCTGCTGCTGCTGCTGAA 600
 DB 541 GCTTGGGAGAACATATGCGGCTTCTGCTGCAACATGCTGCTGCTGCTGCTGCTGAA 600
 QY 601 GGCATGTGCCCCCTAGCCCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
 DB 601 GGCATGTGCCCCCTAGCCCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
 QY 661 TGACTTCAACGCGCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATATGC 720
 DB 661 TGACTTCAACGCGCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATATGC 720
 QY 721 TGACACTAGCGGCTATATGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
 DB 721 TGACACTAGCGGCTATATGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
 QY 781 TGGCGTCACTGTGATGAGCCAGCCAGAGGCGCTGTGAACTGAGAGCTTCTCTGCGCA 840
 DB 781 TGGCGTCACTGTGATGAGCCAGCCAGAGGCGCTGTGAACTGAGAGCTTCTCTGCGCA 840
 QY 841 TGTTCCTGAGAGCTGTGCGCGCACTTTTGCGCCATGTGTGTGTGTGTGTGTGTGTGT 900
 DB 841 TGTTCCTGAGAGCTGTGCGCGCACTTTTGCGCCATGTGTGTGTGTGTGTGTGTGTGT 900
 QY 901 AAGAGGGGGTGGCGAGACCGCGGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 DB 901 AAGAGGGGGTGGCGAGACCGCGGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960

QY 961 TGGGAGATATTTTCCCACTGCAATGTGAAAGAGTGCCTTCCAGCTGCCGAGACACCG 1020
 DB 961 TGGGAGATATTTTCCCACTGCAATGTGAAAGAGTGCCTTCCAGCTGCCGAGACACCG 1020
 QY 1021 GGGAGCCATCGGCTATGGAGAGGCGCAGAGAGGCTGAGAGGCTTGGGCTGGGAGAG 1080
 DB 1021 GGGAGCCATCGGCTATGGAGAGGCGCAGAGAGGCTGAGAGGCTTGGGCTGGGAGAG 1080
 QY 1081 TGTGTAGACCGATGAAAGACCCGAGCTGTAGAGACTCAGAGGCCCATCTTCTTAAGCAG 1140
 DB 1081 TGTGTAGACCGATGAAAGACCCGAGCTGTAGAGACTCAGAGGCCCATCTTCTTAAGCAG 1140
 QY 1141 CCCCCACCTGAGAGAGCCCACTGTTCTCACTTACCTGAGCCCTCAGAGCTCAGCAGA 1200
 DB 1141 CCCCCACCTGAGAGAGCCCACTGTTCTCACTTACCTGAGCCCTCAGAGCTCAGCAGA 1200
 QY 1201 TTTGTCTAAGATAGCAGCAGGATTCAGAGCTTAAGTTGAGGCTGAGATCCAGCTCTCTA 1260
 DB 1201 TTTGTCTAAGATAGCAGCAGGATTCAGAGCTTAAGTTGAGGCTGAGATCCAGCTCTCTA 1260
 QY 1261 ACCCTCAGGCGCAGATGCTTGTCCATGSCACTTTCATGCTCTTGAAGAACTCGAGATGTG 1320
 DB 1261 ACCCTCAGGCGCAGATGCTTGTCCATGSCACTTTCATGCTCTTGAAGAACTCGAGATGTG 1320
 QY 1321 TGAAGGCTATGCTCTGAGACCTGACGGGTTTGTATCTTGAACCTCGGTTACTTTCTG 1380
 DB 1321 TGAAGGCTATGCTCTGAGACCTGACGGGTTTGTATCTTGAACCTCGGTTACTTTCTG 1380
 QY 1381 GGGCCCAAGCTGTGCTGAGACATCTTTCTGCTGAGAGGATTAAGGCTGATTAAT 1440
 DB 1381 GGGCCCAAGCTGTGCTGAGACATCTTTCTGCTGAGAGGATTAAGGCTGATTAAT 1440
 QY 1441 TCTTCTGAGATGACAGTAAACCCAGATGAGAGATAGAGGATCTAGACATGTGCT 1500
 DB 1441 TCTTCTGAGATGACAGTAAACCCAGATGAGAGATAGAGGATCTAGACATGTGCT 1500
 QY 1501 TCTTCCGAAATTTGAGATGATATTTTCAAGGCCCACTTATGATTTCTGATAGCTCTG 1560
 DB 1501 TCTTCCGAAATTTGAGATGATATTTTCAAGGCCCACTTATGATTTCTGATAGCTCTG 1560
 QY 1561 GAGCAGAGGCGAGGAGTTGCAATGTATGATGACCTGCAACATTTGAGAAATTAAGTGA 1620
 DB 1561 GAGCAGAGGCGAGGAGTTGCAATGTATGATGACCTGCAACATTTGAGAAATTAAGTGA 1620
 QY 1621 TCCCTTTGCAACCGCTAGCTAGTATGATTAATTAATTAATTAATTAATTAATTAATTA 1680
 DB 1621 TCCCTTTGCAACCGCTAGCTAGTATGATTAATTAATTAATTAATTAATTAATTAATTA 1680
 QY 1681 GGGTCCCGAGCTAAGGAGCTGCGCTTGGCTTCACTGAGTGAAGAGAGGCGCTGGAGT 1740
 DB 1681 GGGTCCCGAGCTAAGGAGCTGCGCTTGGCTTCACTGAGTGAAGAGAGGCGCTGGAGT 1740
 QY 1741 CTGAACCCAGAGGCTGAGGCGAGGCGAGCTGCGCTAAGTGAAGTGAAGTGAAGTGA 1800
 DB 1741 CTGAACCCAGAGGCTGAGGCGAGGCGAGCTGCGCTAAGTGAAGTGAAGTGAAGTGA 1800
 QY 1801 CAGGCGAGGCGAGCTGATTCAGAGTGCCTCATGAGAGTGAAGGAGCGCTTCCGCGCG 1860
 DB 1801 CAGGCGAGGCGAGCTGATTCAGAGTGCCTCATGAGAGTGAAGGAGCGCTTCCGCGCG 1860
 QY 1861 ATGCAAGGCTGAGGCTGATTCAGAGTGCCTCATGAGAGTGAAGGAGCGCTTCCGCGCG 1920
 DB 1861 ATGCAAGGCTGAGGCTGATTCAGAGTGCCTCATGAGAGTGAAGGAGCGCTTCCGCGCG 1920
 QY 1921 AAAAAAAAAAAAAAAAAA 1939
 DB 1921 AAAAAAAAAAAAAAAAAA 1939

RESULT 4
 AAA46932
 ID AAA46932 standard; cDNA, 1939 BP.
 XX

AC AAA46932;
 XX
 DT 03-OCT-2000 (first entry)
 XX
 DE cDNA encoding novel polypeptide PRO853.
 XX
 KW PRO201, PRO292, PRO327, PRO1265, PRO344, PRO343, PRO347, PRO357,
 KW PRO15, PRO1017, PRO112, PRO509, PRO882, tumour cell,
 KW tumorigenesis; cancer; neoplastic cell growth; cell proliferation; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 128..1261
 FT /tag= a
 XX
 PN MO200037640-A2.
 XX
 PD 29-JUN-2000.
 XX
 PF 16-DEC-1999; 99WO-US30095.
 XX
 PR 22-DEC-1998; 98US-0113296.
 PR 08-MAR-1999; 99WO-US05028.
 PR 02-JUN-1999; 99WO-US12252.
 PR 01-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21090.
 PR 30-NOV-1999; 99WO-US28313.
 PR 30-NOV-1999; 99WO-US28409.
 PR 01-DEC-1999; 99WO-US28301.
 PR 02-DEC-1999; 99WO-US28365.
 XX
 PA (GENTH) GENENTECH INC.
 XX
 PI Botstein D, Goddard A, Gurney AL, Hillian K, Lawrence DA, Roy MA,
 PI Wood WI
 DR WPI; 2000-452188/39.
 DR P-PSDB; AAY93696.
 XX
 PT New anti-polypeptide antibody useful in the treatment and diagnosis of
 PT neoplastic cell growth and proliferation -
 XX
 PS Claim 50; Fig 25; 220bp; English.
 XX
 CC The present sequence encodes a novel human polypeptide. The
 CC specification describes novel polypeptides designated PRO201, PRO292,
 CC PRO327, PRO1265, PRO344, PRO343, PRO357, PRO1017,
 CC PRO112, PRO509, PRO853 and PRO882. These genes are amplified in
 CC the genome of tumour cells. The polypeptides are believed to contribute
 CC to tumorigenesis. The polypeptides are useful target for the
 CC identification of certain cancers, and may act as predictors of the
 CC prognosis of tumour treatment. Antibodies against these polypeptides
 CC are useful in the treatment and diagnosis of neoplastic cell growth
 CC and proliferation in mammals.
 XX
 SQ Sequence 1939 BP; 367 A; 549 C; 624 G; 399 T; 0 other;
 XX
 Query Match 100.0%; Score 1939; DB 21; Length 1939;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCCTCCGCTTCGAGAGCTGAGGCGCGGCGCGCTTCCAGAGCTGTGAGAGGCGGAT 60
 DB 1 CGCCTCCGCTTCGAGAGCTGAGGCGCGGCGCGCTTCCAGAGCTGTGAGAGGCGGAT 60
 QY 61 CGGCAAGCGGCTGAGGCGGATTCAGAGCGGCTGCGGCGCGGAGCGGAGCGCGC 120
 DB 61 CGGCAAGCGGCTGAGGCGGATTCAGAGCGGCTGCGGCGCGGAGCGGAGCGCGC 120
 QY 121 GGGCGGCAATGAGAGGCGCTGCGCGGCGGCGGCTTCTGCTGAGGCGCTTACGTGCTTGT 180
 DB 121 GGGCGGCAATGAGAGGCGCTGCGCGGCGGCGGCTTCTGCTGAGGCGCTTACGTGCTTGT 180

181 CTAATCAACCTGTTGAAAGCCCGCTGTCGGCGGCAATGGGCAACTTCGGGGCCGAC 240
181 CTAATCAACCTGTTGAAAGCCCGCTGTCGGCGGCAATGGGCAACTTCGGGGCCGAC 240
241 GGCCTGTGTCAAGGCGCGCAAGAGGCGCATCGGAAAGATGACGGCGCTGAGCTGGCGCG 300
241 GGCCTGTGTCAAGGCGCGCAAGAGGCGCATCGGAAAGATGACGGCGCTGAGCTGGCGCG 300
301 CCGGAGAGCGCGGTGTGTCTGTGCGCGCAAGCAAGCGCGGAGAGCGCGCTGCTT 360
301 CCGGAGAGCGCGGTGTGTCTGTGCGCGCAAGCAAGCGCGGAGAGCGCGCTGCTT 360
361 CCACTCCCGCGAGAGAGTGGGAAACAATAGGTCATCTTCAATGCGCTTGAATGGCCAG 420
361 CCACTCCCGCGAGAGAGTGGGAAACAATAGGTCATCTTCAATGCGCTTGAATGGCCAG 420
421 TCTGAGCTCGTGTGCGGCGCTTGTGCACTGCTTCTGAGCTCTGAGCCAGCGGTGGACAT 480
421 TCTGAGCTCGTGTGCGGCGCTTGTGCACTGCTTCTGAGCTCTGAGCCAGCGGTGGACAT 480
481 CCTCATCCAAATGCGCGATCAATTCCTGTGCGCGGACCGGTGAGCGCTTTAACTGCT 540
481 CCTCATCCAAATGCGCGATCAATTCCTGTGCGCGGACCGGTGAGCGCTTTAACTGCT 540
541 GCTTGGGGTGAACCATATGCTGCTTCTTCTGAGCAATCTGCTGCTGCTTGGCTGA 600
541 GCTTGGGGTGAACCATATGCTGCTTCTTCTGAGCAATCTGCTGCTGCTTGGCTGA 600
601 GGCATGTGCGCGCTGAGCGCGGTGTGTGTGCTGAGCTGCGCACTGTGCGGAGCGT 660
601 GGCATGTGCGCGCTGAGCGCGGTGTGTGTGCTGAGCTGCGCACTGTGCGGAGCGT 660
661 TGAATTCAAACGCTGGAACCGCGCAAGTGTGCGCTGCGCGAGAGCTGCGGCAATGTC 720
661 TGAATTCAAACGCTGGAACCGCGCAAGTGTGCGCTGCGCGAGAGCTGCGGCAATGTC 720
721 TGAACATAAGCTGCTGATATGATCTGTTTCCCGGAGAGCTGCGCAACCACTGAGGCGAC 780
721 TGAACATAAGCTGCTGATATGATCTGTTTCCCGGAGAGCTGCGCAACCACTGAGGCGAC 780
781 TGGCGTCACTGCTGATGAGCGGCAACCGAGGCGCTGAGCTGAGGCTGCTGCGGCA 840
781 TGGCGTCACTGCTGATGAGCGGCAACCGAGGCGCTGAGCTGAGGCTGCTGCGGCA 840
841 TGTTCCTGTGATGCTGCGGCACTTGTGCGGCAATGTGCTGCTGCTGCGGCAAC 900
841 TGTTCCTGTGATGCTGCGGCACTTGTGCGGCAATGTGCTGCTGCTGCGGCAAC 900
901 AAGAGGGGGTGTGCGGCAACCGCTGTATGTGCTTCAAGAGGCGCATGAGCCCTCAG 960
901 AAGAGGGGGTGTGCGGCAACCGCTGTATGTGCTTCAAGAGGCGCATGAGCCCTCAG 960
961 TGGAGATATTTTGGCAATGCGCATGTGAGAGAGGCGCTCAAGCTGCGCGAGAGAGCGG 1020
961 TGGAGATATTTTGGCAATGCGCATGTGAGAGAGGCGCTCAAGCTGCGCGAGAGAGCGG 1020
1021 GGCAGCGCATCGGCTATGAGAGGCGCAAGAGGCGCTGAGGCTTGGCGCTGGGAGGA 1080
1021 GGCAGCGCATCGGCTATGAGAGGCGCAAGAGGCGCTGAGGCTTGGCGCTGGGAGGA 1080
1081 TGTGTAAACCGCATGAAAGACCCCGCATGTGAGAGCTCAAGAGGCGCATGCTTCTTAAGCAC 1140
1081 TGTGTAAACCGCATGAAAGACCCCGCATGTGAGAGCTCAAGAGGCGCATGCTTCTTAAGCAC 1140
1141 CCCCCACCTGTGAGAGCGCAAGTTCTCAACCTTACCGCAAGCGCTGAGAGCTCAAGCA 1200
1141 CCCCCACCTGTGAGAGCGCAAGTTCTCAACCTTACCGCAAGCGCTGAGAGCTCAAGCA 1200
1201 TTTGTCTAAGATGAGCAACCAATTCAGGCTAAAGTTGAACCTGAGATCAAGCTCTCTTA 1260
1201 TTTGTCTAAGATGAGCAACCAATTCAGGCTAAAGTTGAACCTGAGATCAAGCTCTCTTA 1260

1261 ACCCTCAGGCGCAGAGATGCTTGTGCAATGSCACTTCAAGGTCTTGAAGAACTCGGATGTG 1320
1261 ACCCTCAGGCGCAGAGATGCTTGTGCAATGSCACTTCAAGGTCTTGAAGAACTCGGATGTG 1320
1321 TGAAGCATGCGCTGAGACATGAGCGGCTTGTGATCTTGAACCTCGGTGTTACTTCTG 1380
1321 TGAAGCATGCGCTGAGACATGAGCGGCTTGTGATCTTGAACCTCGGTGTTACTTCTG 1380
1331 TGGGCGATGCGCTGAGACATGAGCGGCTTGTGATCTTGAACCTCGGTGTTACTTCTG 1380
1331 TGGGCGATGCGCTGAGACATGAGCGGCTTGTGATCTTGAACCTCGGTGTTACTTCTG 1380
1381 GGGCCCAAGCTGTGCGCTGAGACATCTTCTTCTGAGTGAAGAAATATGGGTGATTA 1440
1381 GGGCCCAAGCTGTGCGCTGAGACATCTTCTTCTGAGTGAAGAAATATGGGTGATTA 1440
1441 TCTTCTGAGATGACATGAACCCGATGAGAGATGAGGCTATGCTGACCTGTGCT 1500
1441 TCTTCTGAGATGACATGAACCCGATGAGAGATGAGGCTATGCTGACCTGTGCT 1500
1501 TCTGGAATATTTGATATGATATTTTCAAGGCGCGCAACCTTATGATCTGATCAGCTG 1560
1501 TCTGGAATATTTGATATGATATTTTCAAGGCGCGCAACCTTATGATCTGATCAGCTG 1560
1561 GAGCAGAGCGCAGGAGTGTGCAATGTGATGCACTGCGCAATGAGAAATTAAGTGA 1620
1561 GAGCAGAGCGCAGGAGTGTGCAATGTGATGCACTGCGCAATGAGAAATTAAGTGA 1620
1621 TCCCTTGTGCAACCGTCTAGCTAGTATGATTAATTAACCCCATGTTATGAGCGGAATTA 1680
1621 TCCCTTGTGCAACCGTCTAGCTAGTATGATTAATTAACCCCATGTTATGAGCGGAATTA 1680
1681 GGCCTCCGAGCTTAAAGGAGCTGCGCGTGAAGGCTTCAAGTATGAGAGAGGCGCTGGAT 1740
1681 GGCCTCCGAGCTTAAAGGAGCTGCGCGTGAAGGCTTCAAGTATGAGAGAGGCGCTGGAT 1740
1741 CTGAACCCAAAGGCTTGAAGGCGCGCGCACTGCGTAAAGATGAGTGTGCTGAGAGTGA 1800
1741 CTGAACCCAAAGGCTTGAAGGCGCGCGCACTGCGTAAAGATGAGTGTGCTGAGAGTGA 1800
1801 CAGGCGAGGCGAGCTGTGATGAGAGTGTGCGGCAATGAGAGTGAAGGCGCGCTTCCGCGG 1860
1801 CAGGCGAGGCGAGCTGTGATGAGAGTGTGCGGCAATGAGAGTGAAGGCGCGCTTCCGCGG 1860
1861 ATGAGGCGCTGCGGCTGATCTGATCTGAAGCGGCTTGAAGTGAAGCGCTTGAAGCGGCA 1920
1861 ATGAGGCGCTGCGGCTGATCTGATCTGAAGCGGCTTGAAGTGAAGCGCTTGAAGCGGCA 1920
1921 AAAAAAAAAAAAAAAAAA 1939
1921 AAAAAAAAAAAAAAAAAA 1939
RESULT 5
ABX92437
ID ABX92437 standard; cDNA; 1939 BP.
XX
XX ABX92437;
DT 08-MAY-2003 (first entry)
XX
XX cDNA encoding human PRO853 polypeptide.
DE
XX
XX Human; PRO polypeptide; secreted and transmembrane protein;
KW Immune disorder; diabetes; hyper-insulinaemia; hypo-insulinaemia;
KW cardiac insufficiency; nervous system disorder; kidney disorder;
KW bone disorder; cartilage disorder; arthritis; tumour; wound healing;
KW genetic disorder; cytostatic; antidiabetic; antiinflammatory;
KW antifibrotic; anti-tumour; vulnerary; antinaemic; dermatological;
KW cardiac; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX US2002169284-A1.
XX
XX
XX
XX 14-NOV-2002.

```
PF 16-OCT-2001; 2001US-0978697.
XX 07-OCT-1998; 98WO-US21141.
PR 20-NOV-1998; 98WO-US24855.
PR 05-JAN-1999; 99WO-US00106.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 14-MAY-1999; 99WO-US01733.
PR 02-JUN-1999; 99WO-US12252.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 16-DEC-1999; 99WO-US28565.
PR 30-DEC-1999; 99WO-US31243.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US05520.
PR 22-MAR-2001; 2001WO-US09552.
PR 25-MAY-2001; 2001WO-US17092.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 17-OCT-1997; 97US-062250P.
PR 03-NOV-1997; 97US-064249P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066364P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077641P.
PR 12-MAR-1998; 98US-077649P.
PR 13-MAR-1998; 98US-077911P.
PR 20-MAR-1998; 98US-078804P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078910P.
PR 20-MAR-1998; 98US-078936P.
PR 20-MAR-1998; 98US-078939P.
PR 25-MAR-1998; 98US-079294P.
PR 26-MAR-1998; 98US-079656P.
PR 27-MAR-1998; 98US-079663P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079689P.
PR 27-MAR-1998; 98US-079728P.
PR 27-MAR-1998; 98US-079786P.
PR 30-MAR-1998; 98US-079920P.
PR 30-MAR-1998; 98US-079923P.
PR 26-MAY-1981; 81US-0267213.
PR 17-MAR-1998; 98US-0040220.
PR 26-JUN-1998; 98US-0105413.
PR 07-OCT-1998; 98US-0168978.
PR 02-NOV-1998; 98US-0184216.
PR 06-NOV-1998; 98US-0187368.
PR 07-DEC-1998; 98US-0202054.
PR 22-DEC-1998; 98US-0218517.
PR 05-MAR-1999; 99US-0254465.
PR 10-MAR-1999; 99US-0265686.

PR 12-APR-1999; 99US-0284291.
PR 14-MAY-1999; 99US-0311832.
PR 14-MAY-1999; 99US-0380137.
PR 25-AUG-1999; 99US-0380138.
PR 25-AUG-1999; 99US-0380142.
PR 08-NOV-2000; 2000US-0709238.
PR 27-NOV-2000; 2000US-0723749.
PR 20-DEC-2000; 2000US-0747259.
PR 22-MAR-2001; 2001US-0616744.
PR 10-MAY-2001; 2001US-0816920.
PR 10-MAY-2001; 2001US-0854208.
PR 01-JUN-2001; 2001US-0854280.
PR 05-JUN-2001; 2001US-0872035.
PR 14-JUN-2001; 2001US-0874503.
PR 19-JUN-2001; 2001US-0882636.
PR 30-JUN-2001; 2001US-0918585.

(PGTH ) GENENTECH INC.
PI Ashkenazi A, Baker KP, Botstein D, Desnoyers L, Eaton D;
PI Ferrara N, Flyvareff E, Fong S, Gao W, Gerber H, Gertlisen ME;
PI Goddard A, Godowski PJ, Grimaldi JC, Gunney AL, Hillan KJ;
PI Kijavyn IJ, Kuo SS, Napier MA, Pen J, Paothi NF, Roy MA;
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
XX WPI; 2003-288163/28.
DR P-PSDB; ABU61101.
XX
PT Novel secreted and transmembrane polypeptides and polynucleotides
PT encoding them useful for treating cancer, kidney diseases, bone,
PT cartilage disorders and immune deficiencies
XX
PS Claim 2; Fig 74; 459pp; English.
XX
CC The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The
CC PRO polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides are useful for detecting other PRO polypeptides, for
CC linking bioactive molecules to cells expressing PRO polypeptides,
CC for modulating biological activities of cells expressing PRO
CC polypeptides, and for identifying agonists or antagonists. The
CC bioactive molecule may be a toxin, radiolabel or antibody, and causes
CC apoptosis or death of the cell. The PRO polypeptides are useful for
CC treating immune disorders, diabetes or hyper- or hypo-insulinaemia,
CC cardiac insufficiency, nervous system disorders, kidney disorders,
CC bone and cartilage disorders or arthritis, tumours, and wound healing.
CC The polynucleotide sequences encoding PRO polypeptides are useful as
CC hybridisation probes, in chromosome and gene mapping, in the generation
CC of antisense RNA and DNA, in the preparation of PRO polypeptides, for
CC generating transgenic animals or knockout animals, for the genetic
CC analysis of individuals with genetic disorders, and in gene therapy.
CC The present sequence encodes a human PRO polypeptide of the invention.
CC Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at
CC seqdata.uspto.gov/patident.html.
XX
SQ Sequence 1939 BP; 367 A; 549 C; 624 G; 399 T; 0 other;
XX

Query Match 100.0%; Score 1939; DB 25; Length 1939;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCTCGCCTTGGAGAGCTGACGCGCGCGCGCGCTTCCAGGCTGTGACGAGCGGAT 60
DB 1 CGCCTCGCCTTGGAGAGCTGACGCGCGCGCGCGCTTCCAGGCTGTGACGAGCGGAT 60
QY 61 CGGACGCGCTGCGCGCGCTGACGCGCGCGCGCGCTTCCAGGCTGTGACGAGCGGAT 120
DB 61 CGGACGCGCTGCGCGCGCTGACGCGCGCGCGCGCTTCCAGGCTGTGACGAGCGGAT 120
QY 121 GGGCGGATGAGAGCGCGCTGTGCTGGCGCGCGCGCTTCCAGGCTGTGACGAGCGGAT 180
DB 121 GGGCGGATGAGAGCGCGCTGTGCTGGCGCGCGCGCTTCCAGGCTGTGACGAGCGGAT 180
```


Cy		1646	AGTAAATTTACCCCATGTATATGAAAGCGGAATTAGCTCCGAGCTAAGGGACTC	1705
Dz		1574	AGTTAAATTACCCCATTATTATGAAAGCGGAATTAGCTCCGAGCTAAGGGACTC	1633
Oy		1706	AGGCTTCACAAGTAGTAGAGAGAGAGGCCCTTG9GATCTTAACC	1765
Dd		1634	AGGGCTCAcAGTAGTAGAGAGAGAGGGCCTGGAGCTTGACCCAAAGGTTGAGGGC	1693
Oy		1766	GCCGACTGCCTTAAGATGGGTGCTGAGAAGTGAAGTCAAGGCGAGGCTGGTATCGAG	1825
Dd		1694	GCCGACTGCCTTAAGATGGGTGCTGAGAAGTGAAGTCAAGGCGAGGCTGGTATCGAG	1753
Oy		1826	TGCCCCAATGGAGATPAGAGGAGAGCGCCTTCGGGGCGGAGTC	1885
Dd		1754	TGCCCCAATGGAGATPAGAGGAGAGCGCCTTCGGGGCGGAGTC	1813
Oy		1886	TGAAGCCCTCGSAGATAAGAGCGCGTGAACGCCCAAAAAAAAAAAAAAAA	1939
Dd		1814	TGAAGCCCTCGSAGATAAGAGCGCGTGAACGCCCAAAAAAAAAAAAAAAA	1867
RESULT 7				
AAD09318				
ID	AAD09318	standard; cDNA; 1779 BP.		
XX				
AC	AAD09318;			
XX				
DT	12-SEP-2001	(first entry)		
XX				
DE	Human drug metabolizing enzyme (DME-3) cDNA.			
XX				
KM	Human, drug metabolizing enzyme; DME-3; immunosuppressive; gene therapy;			
KM	cystoelastic; autoimmune disorder; inflammatory disorder; atherosclerosis;			
KM	osteoporosis; eye disorder; hepatic tumor; Addison's disease; cretinism;			
KM	rheumatoid arthritis; acquired immune deficiency syndrome; AIDS; anaemia;			
KV	developmental disorder; endocrine disorder; infertility; acromegaly; epilepsy;			
KW	thyrotoxicosis; pancreatic disorder; diabetes mellitus; obesity; adenoma;			
KW	gastrointestinal disorder; nodular hyperplasia; conjunctivitis; glaucoma;			
KM	acinic keratosis; metabolic disorder; dysphagia; anorexia; carcinoma;			
XX	cell proliferative disorder; ss.			
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	252..1112		
FT		/tag= a		
FT	sig_peptide	252..455	"Human drug metabolizing enzyme (DME-3)"	
FT		/tag= b		
FT	mat_peptide	456..1109		
FT		/tag= c		
FT		/product= "Mature drug metabolizing enzyme (DME-3)"		
PN	WO200151638-A2.			
PD	19-JUL-2001.			
PF	12-JAN-2001; 2001WO-US01174.			
PR	14-JAN-2000; 2000US-0176139.			
PR	21-JAN-2000; 2000US-0177443.			
PR	28-JAN-2000; 2000US-0178574.			
PA	(INCYTE GENOMICS INC.			
P1	Yang J, Baughn MR, Burford N, Au-Young J, Lu DM, Reddy R,			
P1	Ring HZ, Hillman JT, Yue H, Azimzal Y, Yao WJ, Gandhi AR,			
P1	Nguyen DB, Tang YT, Lal P, Bandman O,			
DR	WPI, 2001-425874/45.			
DR	P-Psdb; AAE05172.			
PT	Drug metabolizing enzymes and encoding polynucleotides, useful for			

Pt	diagnosing, treating and/or preventing autoimmune, inflammatory, cell
Pt	proliferative, developmental, endocrine, eye, metabolic, and
Pt	gastrointestinal disorders -
Px	
Px	
Px	Claim 5; Page 160; 133pp; English.
Cc	
Cc	The present sequence is human drug metabolizing enzyme (DME-3) cDNA.
Cc	Human DME and its nucleic acid molecule are useful for the diagnosis,
Cc	treatment and prevention of disorders associated with increased or
Cc	decreased expression of DME. Examples of such disorders include,
Cc	autoimmune/inflammatory disorder such as acquired immune deficiency
Cc	syndrome (AIDS), rheumatoid arthritis, osteoporosis; cell proliferative
Cc	disorder such as actinic keratosis, atherosclerosis; developmental
Cc	disorders such as epilepsy, anaemia; endocrine disorder such as
Cc	acromegaly, cretinism, thyrotoxicosis; pancreatic disorder such as
Cc	diabetes mellitus; eye disorder such as conjunctivitis, glaucoma, iritis;
Cc	metabolic disorder such as Addison's disease, obesity; gastrointestinal
Cc	disorder such as anorexia, dysphagia and hepatic tumours including
Cc	nodular hyperplasia, adenomas and carcinomas. DME DNA is useful for
Cc	creating "knockin" humanised animals (pigs) or transgenic animals (mice
Cc	or rats) to model human disease. DME DNA is also in useful in gene
Cc	therapy. DME and its immunogenic fragments are useful for screening
Cc	libraries of compounds in several drug screening assays.
SQ	
SQ	Sequence 1779 BP; 385 A; 475 C; 523 G; 396 T; 0 other:
Query Match	77.2%; Score 1497; DB 22; Length 1779;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 1547; Conservative	0; Mismatches 1; Indels 0; Gaps 0
QY	.371 CAGGAGATGGGAACAAATGAGTCACTTTCAATGCGCTTGGAATTGGCCAGTCGTGGCCTCG 430
DB	222 CAGAGAATGGGAACAATGAGTCACTTTCAATGCGCTTGGAATTGGCCAGTCGTGGCCTCG 281
QY	431 GTGGGGGCGCTTGCACATGCGCTTTCTGAGCTCTGAGGCCAGGATTGGAATCCTATCCAG 490
DB	282 GTGGGGGCGCTTGCACATGCGCTTTCTGAGCTCTGAGGCCAGGATTGGAATCCTATCCAG 341
QY	491 AATGCGCGTATCAGTTCCTGTGAGCGGACCCGTGAGGCGTTTAACCTGCTCCTTGCGGTG 550
DB	342 AATGCGCGTATCAGTTCCTGTGAGCGGACCCGTGAGGCGTTTAACCTGCTCCTTGCGGTG 401
QY	551 AACCATATCGTCCCTTTCTGCTGACACATCTGCTGCTGCTTGCTGAGGCAATGTGCC 610
DB	402 AACCATATCGTCCCTTTCTGCTGACACATCTGCTGCTGCTTGCTGAGGCAATGTGCC 461
QY	611 CCTTAGCGCGGT 670
DB	462 CCTTAGCGCGGT 521
QY	671 CGCCTGACCGCCCAAGT 730
DB	522 CGCCTGACCGCCCAAGT 581
QY	731 CTGGCTATGTACTGTTTTGCCCGGAGCTCCCAACAGCTTTGAGGCGCACTGCGGTACC 790
DB	582 CTGGCTATGTACTGTTTTGCCCGGAGCTCCCAACAGCTTTGAGGCGCACTGCGGTACC 641
QY	791 TGCTATGACGCCAACCCAGGGCGCTGTGAACTCGAGAGCTGTTCTGCGCATGTTCTGGA 850
DB	642 TGCTATGACGCCAACCCAGGGCGCTGTGAACTCGAGAGCTGTTCTGCGCATGTTCTGGA 701
QY	851 TGGCGTGGCGCACTTTTGGCGCCCAATGTGGCTTGGCTGTGGCTCGCGGGGCAANAAGGGGGT 910
DB	702 TGGCGTGGCGCACTTTTGGCGCCCAATGTGGCTTGGCTGTGGCTCGCGGGGCAANAAGGGGGT 761
QY	911 GCCAGACACCCCTGTATTTGTGCTCTCAACAAGGGGCAATGAGCCCTCAGTGGAGATAT 970
DB	762 GCCAGACACCCCTGTATTTGTGCTCTCAACAAGGGGCAATGAGCCCTCAGTGGAGATAT 821
QY	971 TTGGCGAATGCGCATGTGGAAGAGGTGCTTCAGCTGCGCGGAGACAGCCGGGCAAGCCAT 1030
DB	822 TTTCGCACTGCGCATGTGGAAGAGGTGCTTCAGCTGCGCGGAGACAGCCGGGCAAGCCAT 881


```

Db 353 ACCGAGGCGCTCAGAGCTCACACAGATTGTCTTATAGATACGACCGGATTCAGGCTAAG 412
Cy 1236 TTGAGCTGAGATCCAGCTCTCTTAACCTCCAGGCGCAAGATGCTTCCAGATGCTTCAT 1295
Db 413 TTGAGCTGAGATCCAGCTCTCTTAACCTCCAGGCGCAAGATGCTTCCAGATGCTTCAT 472
Cy 1296 GGTCTGTAAGAACTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1355
Db 473 GGTCTGTAAGAACTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 532
Cy 1356 CTGAGCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1415
Db 533 CTGAGCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 592
Cy 1416 GTTGAAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1475
Db 593 GTTGAAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 652
Cy 1476 ATAGGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1535
Db 653 ATAGGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 712
Cy 1536 CCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1595
Db 713 CCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 772
Cy 1596 CCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1655
Db 773 CCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 832
Cy 1656 CCCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1715
Db 833 CCCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 892
Cy 1716 AGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1775
Db 893 AGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 952
Cy 1776 GTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1835
Db 953 GTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1012
Cy 1836 GAGTAAGGGGACGCTCCGCGCGGATGATGATGATGATGATGATGATGATGATGAT 1895
Db 1013 GAGTAAGGGGACGCTCCGCGCGGATGATGATGATGATGATGATGATGATGATGAT 1072
Cy 1896 CGGAATTAAGCGCTTGAACCGCC 1918
Db 1073 CGGAATTAAGCGCTTGAACCGCC 1095

```

RESULT 9
AAF26578
ID AAF26578 standard; DNA; 1140 BP.

XX AAF26578;
XX
XX 27-MAR-2001 (first entry)
XX
XX
XX DNA encoding human secreted protein #32.
XX
XX Secreted protein; gene therapy; vaccine; cancer; leukemia;
XX autoimmune disease; allergy; inflammation; graft rejection;
XX hyperproliferation; cardiovascular; infection; su.
XX
XX Homo sapiens.
XX
XX WO200076531-A1.
XX
XX 21-DEC-2000.
XX
XX 01-JUN-2000; 2000MO-US15137.
XX

```

PR 11-JUN-1999; 99US-0138625.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsu LL GA;
XX
XX WPI; 2001-071148/08.
XX
XX Nucleic acids encoding 47 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX disease and diabetic retinopathy -
XX
XX Claim 1, Page 463; 525pp; English.
XX
XX The present invention relates to 26 secreted human proteins. The
XX proteins may be used in the prevention, diagnosis and treatment of
XX diseases associated with inappropriate polypeptide expression.
XX
XX For example, they may be used in gene therapy or in vaccines.
XX
XX Typical of diseases which are potentially treatable are cancers
XX (including leukemia), autoimmune diseases, allergies, inflammation,
XX graft rejection, hyperproliferation, cardiovascular diseases
XX (particularly critical limb ischemia and coronary disease) and any
XX involving abnormal angiogenesis, neurodegeneration and/or
XX infectious diseases.
XX
XX Sequence 1140 BP; 269 A; 297 C; 330 G; 244 T; 0 other;
XX
XX Query Match 36.0%; Score 698; DB 22; Length 1140;
XX Best Local Similarity 99.5%; Pred. No. 7.3e-262;
XX Matches 1088; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
XX
XX 827 CTGTTCTCGCCCATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 886
XX 13 CTGTTCTCGCCCATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 72
XX
XX 887 GTGCTCCGGGACCAAGAGGGGGGTCGACAGACACCTCTGATGATGATGATGATGATGAT 946
XX 73 GTGCTCCGGGACCAAGAGGGGGGTCGACAGACACCTCTGATGATGATGATGATGATGAT 132
XX
XX 947 ATGAGCCCTCAGTGGGAGATATTTGCAACTGCGATGATGATGATGATGATGATGATGAT 1006
XX 133 ATGAGCCCTCAGTGGGAGATATTTGCAACTGCGATGATGATGATGATGATGATGATGAT 192
XX
XX 1007 GCCGAGACGACCGGGGACGCCATCGCTATGATGATGATGATGATGATGATGATGATGAT 1066
XX 193 GCCGAGACGACCGGGGACGCCATCGCTATGATGATGATGATGATGATGATGATGATGAT 252
XX
XX 1067 GGGCTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126
XX 253 GGGCTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 312
XX
XX 1127 TCTTCTCAACACCCCGCCAGCTGAGAGGCGACAGTTCTCAACCTTACCCAGCCCT 1186
XX 313 TCTTCTCAACACCCCGCCAGCTGAGAGGCGACAGTTCTCAACCTTACCCAGCCCT 372
XX
XX 1187 CAGAGCTCAGCAATTTGCTTAAGATGACAGACCGAATTCAGGCTTAAGTTGAGGCTGAG 1246
XX 373 CAGAGCTCAGCAATTTGCTTAAGATGACAGACCGAATTCAGGCTTAAGTTGAGGCTGAG 432
XX
XX 1247 ATCCAGCTCTCTTAACCTCAGGCGGATGATGATGATGATGATGATGATGATGATGAT 1306
XX 433 ATCCAGCTCTCTTAACCTCAGGCGGATGATGATGATGATGATGATGATGATGATGAT 492
XX
XX 1307 ACCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366
XX 493 ACCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 551
XX
XX 1367 TGGTTACTTTCTGAGG-CCCGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1425
XX 552 TGGTTACTTTCTGAGG-CCCGAGCTGATGATGATGATGATGATGATGATGATGATGAT 611
XX
XX 1426 TATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1485

```

Db 612 TAATGGGATGATTTCTTCTCTGAGAGTACACGACCCCAAGTGGAGATAGGGGAT 671
 Qy 1486 GCTAGACACTGTGCTTCTCGGAAATTTGATGATGATTTTCAAGCCCACTATTTGA 1545
 Db 672 GCTAGACACTGTGCTTCTCGGAAATTTGATGATGATTTTCAAGCCCACTATTTGA 731
 Qy 1546 TTCTGATCAGCTCTGAGACAGGAGGAGATTGCAATGATGACATGCCAATTGA 1605
 Db 732 TTCTGATCAGCTCTGAGACAGGAGGAGATTGCAATGATGACATGCCAATTGA 791
 Qy 1606 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1665
 Db 792 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 851
 Qy 1666 AATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1725
 Db 852 AATGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 911
 Qy 1726 AGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1785
 Db 912 AGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 971
 Qy 1786 TGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1845
 Db 972 TGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1031
 Qy 1846 AGGCTTCCGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1905
 Db 1032 AGGCTTCCGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
 Qy 1906 CGCGTGAACCGCC 1918
 Db 1092 CGCGTGAACCGCC 1104

RESULT 10

AB267229 standard; cDNA, 1140 BP.

AB267229;

26-MAR-2003 (first entry)

Human secreted protein encoding cDNA SEQ ID NO 349.

Human; secreted protein; neurotrophic; neuroprotective; cytosolic;
 virulence; dermatological; immunosuppressive; anti-inflammatory; anti-HIV;
 vulnerability; antibacterial; antiparkinsonian; antisticking; antianemic;
 antitubercular; cancer; antirheumatic; hepatocellular; cerebroprotective;
 antiinflammatory; antidiabetic; antidiabetic; antitumor; anticonvulsant;
 antifungal; antiparasitic; cardiac; immune disorder; infection; vaccine;
 cardiovascular disorder; neurological disease; nephroprotective;
 gene therapy; gene; de.

Homo sapiens.

WO20027186-A2.

03-OCT-2002.

26-MAR-2002; 2002WO-US09188.

27-MAR-2001; 2001US-278650P.

12-SEP-2001; 2001US-0950082.

12-SEP-2001; 2001US-0950083.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI, 2003-040583/03.

P-PsDB; ABP99808.

FT New human secreted proteins encoded by genes contained in cDNA clones
 (e.g. HGAC19), useful for preventing, treating or diagnosing e.g.
 AIDS, multiple sclerosis, herpes virus, leukemia, tick-borne
 encephalitis or West Nile fever
 Claim 7, Page 1350-1351, 2423pp, English.
 The invention relates to novel human genes (AB266891-AB268209) and the
 encoded secreted proteins (ABP99470-ABP99812) useful for preventing,
 treating or ameliorating medical conditions e.g. by protein or gene
 therapy. The genes are isolated from a range of human tissues disclosed
 in the specification. The nucleic acids, proteins, antibodies and
 (ant)agonists are useful in the diagnosis, treatment and prevention of:
 (a) cancer, e.g. breast and ovarian cancer and other cancers of the
 adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 lung or urogenital; (b) immune disorders e.g. Addison's disease,
 allergies, autoimmune haemolytic anemia, autoimmune thyroiditis,
 diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 arthritis and ulcerative colitis; (c) cardiovascular disorders such as
 myocardial ischemias; (d) wound healing; (e) neurological diseases e.g.
 cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
 bacterial, fungal and parasitic infections.

Sequence 1140 BP, 269 A, 297 C, 330 G, 244 T, 0 other;

Query Match 36.04; Score 698; DB 25; Length 1140;

Best Local Similarity 99.58; Pred. No. 7.38-262;

Matches 1086; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy 827 CTGTTCTGCGGCAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 886
 Db 13 CTGTTCTGCGGCAATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 72
 Qy 887 GTGCTCCGAGGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 946
 Db 73 GTGCTCCGAGGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 132
 Qy 947 ATCGAGCCCTCAGTGGAGATATTTTCCCACTGCAATGCAATGCAATGCAATGCAATGCAAT 1006
 Db 133 ATCGAGCCCTCAGTGGAGATATTTTCCCACTGCAATGCAATGCAATGCAATGCAATGCAAT 192
 Qy 1007 GCGGAGACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1066
 Db 193 GCGGAGACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 252
 Qy 1067 GGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1126
 Db 253 GGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 312
 Qy 1127 TCTTCTGAGCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAAC 1186
 Db 313 TCTTCTGAGCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAAC 372
 Qy 1187 CAGAGCTCAGCAATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246
 Db 373 CAGAGCTCAGCAATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
 Qy 1247 ATCAGGCTCCTAACCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1306
 Db 433 ATCAGGCTCCTAACCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 492
 Qy 1307 ACTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1366
 Db 493 ACTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 551
 Qy 1367 TGCTTACTTTCTGGGG-CCCCAGCTGTGCTTCTGATGATGATGATGATGATGATGATGAT 1425
 Db 552 TGCTTACTTTCTGGGGCCCCAGCTGTGCTTCTGATGATGATGATGATGATGATGATGAT 611
 Qy 1426 TAATGGGATGATTTCTTCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1485
 Db 612 TAATGGGATGATTTCTTCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 671

XX 09-APR-2001 (first entry)
 XX Novel human polynucleotide, SEQ ID NO: 2250.
 XX
 XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;
 XX breast cancer; lung cancer; cancer detection; 89.
 XX Homo sapiens.
 XX WO200102568-A2.
 XX
 XX 11-JAN-2001.
 XX
 XX 30-JUN-2000; 2000WO-US18374.
 XX
 XX 02-JUL-1999; 99US-0142310.
 XX 02-JUL-1999; 99US-0142311.
 XX
 XX (CHIR) CHIRON CORP.
 XX (HYSB-) HYSBQ INC.
 XX
 XX Williams LT, Sacobedo J, Innis MA, Garcia PD, Klinger J, Kassam A,
 XX Reinhard C, Randazzo F, Kennedy GC, Pot D, Jamson G, Drmanac R,
 XX Chtenjakov R, Drmanac S, Dickson M, Labat I, Leishkowitz D,
 XX Kita D, Garcia V, Jones LW, Strache-Crain B,
 XX WPI; 2001-091805/10.
 XX
 XX Library of polynucleotides for diagnosing a cancerous state of a
 XX mammalian cell and detecting cancer, particularly of the colon or
 XX prostate, comprises 3351 human polynucleotide sequences -
 XX
 XX Claim 9; Page 873; 1046pp; English.
 XX
 XX The present sequence is one of 3351 sequences in a library of human
 XX polynucleotides. The library is used to detect differentially expressed
 XX genes correlated with a cancerous state of a mammalian cell and can
 XX detect colon, prostate, breast and lung cancer. The library can be used
 XX to produce probes for detection of mRNA and to produce additional copies
 XX of the polynucleotides. The probes can be used for chromosome mapping of
 XX the polynucleotide and for detection of transcription levels. Ribozymes
 XX or antisense oligonucleotides can be generated. The polynucleotides and
 XX their gene products are used as genetic or biochemical markers (e.g. in
 XX blood or tissues) that will detect the earliest changes along the
 XX carcinogenesis pathway and/or monitor the efficacy of therapies and
 XX preventive interventions. The polynucleotides, polypeptides and
 XX antibodies against them can be used in pharmaceutical compositions to
 XX treat the cancers and proliferative disorders such as neoplasia,
 XX dysplasia and hyperplasia.
 XX
 XX Sequence 397 BP; 74 A; 116 C; 125 G; 82 T; 0 other;
 XX
 XX Query March 10.3%; Score 200; DB 22; Length 397;
 XX Best Local Similarity 100.0%; Pred. No. 3.7e-68;
 XX Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 716 TATCTGCACTAAGTGGCTAATGTAATGTTTCCCGGAGCTGCGCAACAGCTTGG 775
 XX 31 TATCTGCACTAAGTGGCTAATGTAATGTTTCCCGGAGCTGCGCAACAGCTTGG 90
 XX
 XX 776 GCCACTGCGCTACCTGCTATGACGCCACCGAGGCGCTTGTAACTGGAAGCTGTTCTG 835
 XX 91 GCCACTGCGCTACCTGCTATGACGCCACCGAGGCGCTTGTAACTGGAAGCTGTTCTG 150
 XX
 XX 836 CGCATATGCTCTGATGAGTGGCGCGCACTTTGGCGCCCATTTGGCTGGCTGCTCGG 895
 XX 151 CGCATATGCTCTGATGAGTGGCGCGCACTTTGGCGCCCATTTGGCTGGCTGCTCGG 210
 XX
 XX 896 GCACCAAGAGGGGTGCGCA 915
 XX 211 GCACCAAGAGGGGTGCGCA 230

RESULT 13
 AAT19666
 ID AAT19666 standard; cDNA to mRNA; 186 BP.
 XX
 XX AAT19666;
 XX
 XX 28-JUN-1996 (first entry)
 XX
 XX Human gene signature HUMGS00736.
 XX
 XX DE
 XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 XX human; cloning; mapping; non-biased library; diagnosis; detection;
 XX cell typing; abnormal cell function; 88.
 XX Homo sapiens.
 XX
 XX WO9514772-A1.
 XX
 XX 01-JUN-1995.
 XX
 XX 11-NOV-1994; 94WO-JF01916.
 XX
 XX 12-NOV-1993; 93JP-0355504.
 XX
 XX (MATS/) MATSUBARA K.
 XX (OKUB/) OKUBO K.
 XX
 XX Matsubara K, Okubo K;
 XX
 XX WPI; 1995-206931/27.
 XX
 XX Identifying gene signatures in 3'-directed human cDNA library - e.g.
 XX for diagnosis of abnormal cell function, by preparing cDNA that
 XX reflects relative abundance of corresp. mRNA in specific human
 XX tissues
 XX
 XX Claim 1; Page 445; 2245pp; Japanese.
 XX
 XX A single-stranded DNA (or its complementary strand or the corresp.
 XX double-stranded DNA) which comprises one of the 7837 "GS" sequences
 XX given in AAT19001-T26837 and which is able to hybridize to part of
 XX human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 XX sequences were obtained from 3'-directed cDNA libraries prepared
 XX from various human tissues; synthesis of cDNA was initiated from the
 XX 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 XX untranslated sequence is unique to a particular mRNA species, almost
 XX all the 3'-oriented cDNAs hybridize with specific mRNAs. Each library
 XX is constructed so as to reflect accurately the relative abundance of
 XX different mRNAs in the particular tissue from which it was derived.
 XX The appearance frequency of a given GS in a cDNA library can be
 XX determined (esp. using primers and probes derived from the GS
 XX sequences) as a means of diagnosing abnormal cell function or for
 XX recognising different cell types.
 XX
 XX Sequence 186 BP; 42 A; 42 C; 67 G; 31 T; 4 other;
 XX
 XX Query March 4.7%; Score 92; DB 16; Length 186;
 XX Best Local Similarity 100.0%; Pred. No. 4.1e-26;
 XX Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1744 AACCCAAAGGCTGAGCGCCAGCGGCTGCTAAGATGGTGTGAGAGTGAAGTCAAG 1803
 XX 8 AACCCAAAGGCTGAGCGCCAGCGGCTGCTAAGATGGTGTGAGAGTGAAGTCAAG 67
 XX
 XX 1804 GGCAGGGCAGCTGTATGAGAGTGGCCCATGG 1835
 XX 68 GGCAGGGCAGCTGTATGAGAGTGGCCCATGG 99
 XX
 XX RESULT 14
 XX ABK64203/c
 XX ID ABK64203 standard; DNA; 175 BP.

